

DOCUMENT RESUME

ED 250 327

TM 840 533

AUTHOR Rule, Judy Gayle
TITLE Effects of Multigrade Grouping on Elementary Student Achievement in Reading and Mathematics. RR-48.
INSTITUTION Mesa Public Schools, AZ. Dept. of Research and Evaluation.
PUB DATE May 83
NOTE 143p.; This paper received the 1984 American Educational Research Association Division H award for Best Report of Instructional Research.
PUB TYPE Reports - Evaluative/Feasibility (142)
EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Ability Grouping; Achievement Gains; Achievement Tests; Analysis of Variance; Elementary Education; *Grouping (Instructional Purposes); *Mathematics Achievement; *Multigraded Classes; *Program Evaluation; *Reading Achievement
IDENTIFIERS California Achievement Tests; *Mesa Public Schools AZ

ABSTRACT

This study investigated the effects of multigrade classes on student achievement in the Mesa Public Schools in Reading and Mathematics. Student achievement in differing abilities groups and levels was also analyzed. The District formed multigrade classes from adjacent grade levels to reduce class loads and number of teachers. Students retained their grade-level assignments and maintained their grade-specific curricula. The sample included 3,360 third through sixth graders in three groups: multigrade classes; single-grade classes from schools with multigrade classes; and single-grade classes from schools without multigrade classes. The 1981 and 1982 California Achievement Tests results were analyzed. Results indicated multigrade classes have no detrimental effects on reading or mathematics achievement with one exception: the mathematics achievement of average students. Using multigrade classes with only average/high students and identifying appropriate instructional techniques was recommended. An eight-page bibliography is included. (BS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED250327

EFFECTS OF MULTIGRADE GROUPING ON
ELEMENTARY STUDENT ACHIEVEMENT
IN READING AND MATHEMATICS

Judy Gayle Rule

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. S. De Grasse

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

DEPARTMENT
OF
RESEARCH & EVALUATION

MESA PUBLIC SCHOOLS



TM 840 533

EFFECTS OF MULTIGRADE GROUPING ON
ELEMENTARY STUDENT ACHIEVEMENT
IN READING AND MATHEMATICS

Judy Gayle Rule

Dr. George N. Smith
SUPERINTENDENT

Dr. James K. Zaharis
DEPUTY SUPERINTENDENT
EDUCATIONAL SERVICES

Dr. James S. DeGracie
DIRECTOR
DEPARTMENT OF RESEARCH AND EVALUATION

May 1983
MESA PUBLIC SCHOOLS

RR-48

FOREWORD

The major function of the Mesa Public Schools Department of Research and Evaluation is to provide information to decision makers. This data should be timely, valid, reliable, accurate, and informative. It should be of sufficient depth, yet concise enough to speak squarely to the issue at hand.

The purpose of this report by the Department of Research and Evaluation was to answer questions raised concerning the use of multigrade classes in the Mesa Schools. It is felt that the information gathered and the resulting analysis has yielded answers to a number of the questions posed.

The major effort in the development of this report was the responsibility of Research and Evaluation, under the direction of Dr. James DeGracie. Ms. Sarah Peterson contributed greatly in the data analysis. Ms. Judy Hill accomplished extensive typing and editing of this report. Mesa Public Schools Data Processing Department, particularly Scott Bair and Bill Neiggemann, provided a significant effort in the sample selection.

Research and Evaluation is a department answering to the Deputy Superintendent, Dr. James Zaharis.

Effects of Multigrade Grouping on
Elementary Student Achievement
in Reading and Mathematics

Executive Summary

INTRODUCTION

Discussion

The Department of Research and Evaluation was asked by the Superintendency and, in particular, the Assistant Superintendent for Educational Services at the elementary level to investigate the impact of multigrade classrooms on student achievement. The Mesa Public Schools has a history of using multigrade/multiage classrooms as a means of delivering educational goals and objectives to students. This history can be traced back to the original one room schoolhouse in the city of Mesa. Over the years the schools in Mesa followed the national trend of having one grade serve students at a given age.

In the 1960s and early 1970s, Mesa, based on research findings and philosophical considerations, again experimented with multigrade/multiage placement in a number of elementary schools in the district. Although these experiments met with reasonable success, multigrade/multiage grouping was not adopted as a district philosophy. At the same time these studies were being conducted, the press of finances and class load dictated the use of multigrade classes. The Mesa Public Schools, as many school districts throughout the nation, were forced into using multigrade classes

more and more in an attempt to both economize and equalize class loads.

These considerations and evolving utilization of multigrade classes formed the basis for the Assistant Superintendent for Educational Services at the elementary level to request a study to be done by the Department of Research and Evaluation. The major question being asked was whether or not placement into multigrade classes had any detrimental effect on student achievement. In pursuing the answer to this question, it was felt that a thorough research study should be conducted; that is, a comprehensive literature review should accompany the investigation of student performance. It is felt that this report does just that: it merges a comprehensive review of the research literature with an indepth study of the academic performance of Mesa elementary students in multigrade classes.

Background

Most of the fourteen investigations of multigrade classes from the 1930s to the present indicated no significant differences between achievement gains of students in multigrade and single-grade classes. Four studies evidenced results that were mixed, one study favored single-grade classes, and the remaining nine studies showed no significant differences. All four of the primary-grade studies (Harvey, 1974; MacDonald, 1974; Adair, 1977; Lincoln, 1981) showed no significant differences between single-grade and multigrade classes. Of the three studies limited to intermediate students, two (Dryer, 1949; Adams, 1953) showed no significant differences and one (Foshay, 1948) reported results

favoring single-grade classes.

Of the seven multigrade studies that included both primary and intermediate students, three (Knight, 1938; Chace, 1961; Way, 1969) reported statistically nonsignificant results and four showed mixed results. In the Rehwoldt and Hamilton study (1957), reading achievement favored multigrade classes at first, third, and fifth grades, and language results showed no significant differences. The Finley and Thompson (1963) study had arithmetic results that favored multigrade classes at third grade but favored single-grade boys at fifth grade; other results were not statistically significant. Yerry's (1964) study showed no significant differences between single-grade and multigrade classes at grades two, three, and six; significant differences in grade one favoring multigrade classes in arithmetic, language, and total achievement; and at grade five significant differences favoring multigrade classes in reading, arithmetic, and total achievement. The Milburn (1981) results favored multigrade classes in vocabulary but found no significant differences on other tests.

Only two multigrade studies (Yerry, 1964; Milburn, 1981) considered the achievement of lower- and upper-grade placement in the multigrade classes, and both reported mixed results. Yerry's study favored upper-grade students at primary grades but found no differences at intermediate grades. Milburn found that students in the lower grade outscored their single-grade counterparts, while students in the upper grade showed no differences. None of the studies considered the criterion for

placing students in the multigrade classes, e.g., ability grouping.

Because ability grouping was used as the major criterion for placing students in the multigrade classes, the research on ability grouping was reviewed. Of seven research reviews of ability grouping, only one early review (Cornell, 1936) favored ability grouping. More recent reviews (Wilhelms, 1961; Eckstrom, 1961; Eash, 1961; National Education Association, 1968; Esposito, 1973; Findlay and Bryan, 1970) suggested that ability grouping by itself did not increase achievement. In fact, it was teaching methods, materials, and other variables that had greater effects than the grouping procedure. Of eight recent studies, five (Morgenstern, 1963; Balow, 1962; Lunn, 1970; Borg, 1964; Degrow, 1963) showed no significant differences in achievement. A study by Berkun (1966) favored homogeneous grouping, and a study by Koontz (1961) favored heterogeneous grouping. A comprehensive study by Goldberg and Passow (1966) reported mixed results: only in science did the presence of gifted students increase the results of other students. In math the presence of slow students was consistently related to increased achievement of all other students, and in all academic areas greater gains were shown for the most heterogeneous classes. Thus, though most research showed no significant differences between heterogeneous and homogeneous groupings, the results remained inconclusive (Bolvin, 1982).

No matter what the results of research on a national level, the Mesa Public Schools wanted to insure that no detrimental effect was caused by the use of multigrade classrooms in its

schools. Therefore, this study was designed and conducted to ascertain the results of the use of multigrade classes in the Mesa Public Schools.

METHODS

The Mesa Public School District forms multigrade classes to reduce class loads at two grade levels within a single school and in so doing reduces the number of teachers that otherwise would be required for single-grade classes at those grade levels. The multigrade classes are combined from two adjacent levels, e.g., third and fourth grades, or fourth and fifth grades. Students from the two grades are placed with one teacher but retain their grade-level assignments and maintain their grade-specific curriculums.

The purpose of this study was to determine the effects of multigrade classes on student achievement in the Mesa Schools in reading and mathematics. In addition, the achievement of students in differing ability groups and levels was analyzed. The study analyzed the achievement differences in the following areas.

1. Grouping patterns: Group 1 - multigrade classes,
Group 2 - single-grade classes in multigrade schools,
Group 3 - single-grade classes not in multigrade schools;

2. Placement: Placement 1 - multigrade classes with high-achieving students,

Placement 2 - multigrade classes with average-to high-achieving students,

Placement 3 - multigrade classes with average-achieving students;

3. Levels: Upper - a fourth-grade student in a third/fourth multigrade class,

Lower - a fourth-grade student in a fourth/fifth multigrade class; and

4. Grades: third, fourth, fifth, and sixth.

The sample included 3,360 students in Mesa Public Schools grades three through six in the 1981-82 school year who had 1981 and 1982 California Achievement Tests (CAT) scores available. Students were selected from multigrade classes (Group 1), single-grade classes at the fifteen schools having both single-grade and multigrade classes (Group 2), single-grade classes at the seventeen schools in the district that had no multigrade classes (Group 3).

The one-way analysis of variance using a posttest only design was used. If significant differences were found, Fisher's LSD was used to compare the means.

RESULTS

The results of the twelve one-way analyses of variance tests on reading achievement are presented in Table 1. For the twelve analyses comparing groups on reading achievement, only one analysis

showed a significant difference: high-achieving students in grade four in multigrade classes scored significantly higher than high-achieving fourth-grade students in single-grade classes. At other grades and placements, no significant differences among the groups were found.

Considering the means from the twelve analyses presented in Table 1, eight of the twelve comparisons favored multigrade classes. That is, ignoring statistical significance, the mean scores in reading for students in multigrade classes were higher than those for students in single-grade classes for eight of the twelve comparisons. All four comparisons for Placement 2, the Average/High students favored the multigrade classes. Of the four comparisons that did not favor multigrade classes, three were for Placement 3, the Average students. For all four of the comparison not favoring multigrade, the highest mean was obtained by students in the single-grade classes not in multigrade schools.

For the twelve analyses comparing groups on mathematics achievement (Table 2), five analyses showed significant differences: high-achieving third-graders in single-grade classes scored significantly higher than high-achieving third-graders in multigrade classes; average-achieving third-, fifth-, and sixth-graders in single-grade classes not in multigrade schools outscored average third-, fifth-, and sixth-graders in multigrade classes; average/high-achieving sixth-graders in multigrade classes scored higher than average/high achieving sixth-graders in single-grade classes.

Considering the means from the twelve analyses presented in

Table 1

**Summary of Group Within Placement and Grade on the 1982
California Achievement Tests--Total Reading**

Placement and Group*	Grade							
	Three		Four		Five		Six	
	N	Mean	N	Mean	N	Mean	N	Mean
Placement 1: High								
Group 1 ..	71	76.2	36	78.7*	13	83.7	10	83.7
Group 2 ..	101	72.1	28	69.9*	10	79.5	8	81.5
Group 3 ..	120	74.8	48	71.6*	10	86.1	11	83.3
Placement 2: Average/High								
Group 1 ..	29	67.8	34	67.3	78	67.5	30	66.4
Group 2 ..	157	65.7	102	62.3	358	64.0	162	63.5
Group 3 ..	84	67.4	226	65.1	316	65.8	216	65.9
Placement 3: Average								
Group 1 ..	12	51.3	15	57.3	64	52.9	34	54.3
Group 2 ..	13	54.3	30	61.3	139	53.2	145	52.1
Group 3 ..	51	56.7	56	61.3	320	56.7	223	52.9

*Significant at $p < .05$.

○ Indicates highest mean

Group 1 Multigrade classes

Group 2 Single-grade classes in multigrade schools

Group 3 Single-grade classes not in multigrade schools

Table 2

Summary of Group Within Placement and Grade on the 1982
California Achievement Tests--Total Mathematics

Placement and Group	Grade							
	Three		Four		Five		Six	
	N	Mean	N	Mean	N	Mean	N	Mean
Placement 1: High								
Group 1 ..	70	70.5*	36	79.3	13	85.5	10	84.4
Group 2 ..	46	71.6*	18	73.7	7	76.4	4	90.3
Group 3 ..	72	76.2*	40	76.6	5	88.6	4	89.3
Placement 2: Average/High								
Group 1 ..	31	63.4	33	68.6	73	64.8	30	73.3*
Group 2 ..	137	64.9	95	62.7	348	61.5	154	65.9*
Group 3 ..	67	63.6	204	65.9	288	64.1	209	67.8*
Placement 3: Average								
Group 1 ..	8	57.5*	11	55.0	60	50.3*	33	50.8*
Group 2 ..	9	49.4*	22	58.9	125	51.3*	139	54.0*
Group 3 ..	25	64.6*	46	59.2	256	55.3*	204	60.4*

*Significant at $p < .05$.

○ Indicates highest mean

Group 1 Multigrade classes

Group 2 Single-grade classes in multigrade schools

Group 3 Single-grade classes not in multigrade schools

Table 2, four of the twelve favored multigrade classes. That is, ignoring statistical significance, the mean scores in mathematics for students in multigrade classes were higher than those for students in single-grade classes for four of the twelve comparisons. Three of the four comparisons for Placement 2, the Average/High students favored the multigrade classes. All four comparisons for Placement 3, the Average students favored the single-grade classes not in multigrade schools.

A comparison between student reading scores from Level 1 and Level 2 is summarized in Table 3. The means for each Level by Placement and Grade are shown. For example, the fifth grade students in the fifth/sixth class had a mean of 67.3; the fifth grade students in the fourth/fifth combination had a mean of 67.6. Of the nine possible comparisons of reading scores, three grades by three Placements, that could be made between Level 1 and Level 2, five had sufficient sample size for valid comparisons. No significant differences were found for reading.

Considering the means from the five analyses presented in Table 3, three favored Level 2. That is, ignoring statistical significance, the mean scores in reading for students in multigrade classes did not tend to favor students in either the upper grade or lower grade.

Comparisons between Level 1 and Level 2 student scores in mathematics are presented in Table 4. Of the nine possible comparisons of mathematic scores that could be made, five had sufficient sample size for valid comparisons. Levels 1 and 2 were significantly different for Grade five, Placement three, with the

Table 3
Summary of Level on the 1982 California Achievement
Tests--Total Reading

Placement and Level	Grade					
	Three		Four		Five	
	N	Mean	N	Mean	N	Mean
Placement 1: High						
Level 1	30	76.0	not enough cases		not enough cases	
Level 2	41	76.4				
Placement 2: Average/High						
Level 1	12	70.3	not enough cases		37	67.3
Level 2	17	66.1			41	67.6
Placement 3: Average						
Level 1	not enough cases		7	51.4	14	50.0
Level 2			8	62.5	16	41.6

No significance found at $p < .05$

○ Indicates highest mean

Level 1 Lower grade in multigrade class

Level 2 Upper grade in multigrade class

Table 4
Summary of Level on the 1982 California Achievement
Tests--Total Mathematics

Placement and Level	Grade					
	Three		Four		Five	
	N	Mean	N	Mean	N	Mean
Placement 1: High						
Level 1	30	72.5	not enough cases		not enough cases	
Level 2	25	66.9				
Placement 2: Average/High						
Level 1	8	67.1	not enough cases		33	66.4
Level 2	6	63.2			40	63.5
Placement 3: Average						
Level 1	not enough cases		6	55.3	15	47.5*
Level 2			5	54.6	15	39.1*

*Significant at $p < .05$.

○ Indicates highest mean

Level 1 Lower grade in multigrade combination

Level 2 Upper grade in multigrade combination

means favoring the lower level.

Considering the means from the five analyses presented in Table 4, all five favored Level 1, the lower grade. That is, ignoring statistical significance, the mean scores in mathematics for students in multigrade classes favored those students in the lower grade.

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS
--

1. Multigrade classes did not appear to affect reading achievement negatively. Performance of students in multigrade classes was higher for Placement 1, high achieving students, and Placement 2, average/high achieving students. The performance of students in Placement 3, average students, seemed to be poorer for students in multigrade classes; the achievement of average students was higher in Group 3, single grade classes not in multigrade schools.

2. Multigrade classes appeared to affect mathematics achievement scores somewhat negatively. In this case those students in single grade classes not in multigrade schools appeared to do better than students in multigrade classes, particularly average students. Students in multigrade classes seemed to do better in Placement 2, students in the average to high category. This was especially true for grades 4, 5, and 6.

3. For multigrade classes, Level did not appear to affect

reading achievement. That is, those students who were in the lower grade in multigrade classes versus those students that were in the upper grade in multigrade classes did not perform significantly different. For example, fourth grade student performance in a 4/5 class was not significantly different from fourth grade students' performance in a 3/4 class.

4. For multigrade classes, student performance was not significantly different between Level 1 and Level 2 in mathematics achievement except for Placement 3, average students, at the fifth grade. In this case performance favored Level 1, students in the lower grade in a multigrade class. Overall, in mathematics performance, students in Level 1 seem to perform somewhat better than students in Level 2. That is, students in the lower grade seem to perform better than students in the upper grade in a multigrade class.

From the results of this study it is concluded that no detrimental effect in reading or mathematics achievement occurs with the use of multigrade classes with one exception. This exception is in mathematics achievement of average students in multigrade classes. It is, therefore, recommended that multigrade classes be used only with average/high students. It is also recommended that specific instructional techniques be identified and analyzed for use in multigrade classes. This is especially true in the area of mathematics.

TABLE OF CONTENTS

	Page
LIST OF TABLES	xviii
 CHAPTER	
1. Introduction	1
Discussion	1
Background	2
Statement of the Problem	4
Definition of Terms	6
Limitations	8
Purpose of the Research	9
Summary	12
2. Review of the Research	13
History of Grouping/Grading	14
Research on Nongrading Versus Single-Grading	22
Research on Multigrading Versus Single-Grading	29
Ability Grouping	40
Summary	52
3. Methodology	58
Background	58
Population and Sample	61
Instrumentation	72
Design	73
Analysis of Data	74

CHAPTER	Page
4. Findings	75
5. Summary, Discussion, Conclusions, and Recommendations	99
Summary	99
Discussion	103
Conclusions	107
Implications for Practitioners	108
Recommendations for Researchers	109
BIBLIOGRAPHY	111

LIST OF TABLES

Table	Page
1. Summary of Group Within Placement and Grade on the 1981 California Achievement Tests-- Total Reading	65
2. Summary of Group Within Placement and Grade on the 1981 California Achievement Tests-- Total Mathematics	66
3. Summary of Level Within Placement and Group on the 1981 California Achievement Tests-- Total Reading	67
4. Summary of Level Within Placement and Group on the 1981 California Achievement Tests-- Total Mathematics	69
5. Summary of Analysis of Variance of Reading Achievement by Group, Grade, and Placement	70
6. Summary of Analysis of Variance of Mathematics Achievement by Group, Grade, and Placement	71
7. Summary of Group Within Grade 3 and Placement 1 on the 1982 California Achievement Tests--Total Reading	76
8. Summary of Group Within Grade 3 and Placement 2 on the 1982 California Achievement Tests--Total Reading	76
9. Summary of Group Within Grade 3 and Placement 3 on the 1982 California Achievement Tests--Total Reading	76
10. Summary of Group Within Grade 4 and Placement 1 on the 1982 California Achievement Tests--Total Reading	77

Table

Page

11.	Summary of Group Within Grade 4 and Placement 2 on the 1982 California Achievement Tests--Total Reading	77
12.	Summary of Group Within Grade 4 and Placement 3 on the 1982 California Achievement Tests--Total Reading	77
13.	Summary of Group Within Grade 5 and Placement 1 on the 1982 California Achievement Tests--Total Reading	78
14.	Summary of Group Within Grade 5 and Placement 2 on the 1982 California Achievement Tests--Total Reading	78
15.	Summary of Group Within Grade 5 and Placement 3 on the 1982 California Achievement Tests--Total Reading	78
16.	Summary of Group Within Grade 6 and Placement 1 on the 1982 California Achievement Tests--Total Reading	79
17.	Summary of Group Within Grade 6 and Placement 2 on the 1982 California Achievement Tests--Total Reading	79
18.	Summary of Group Within Grade 6 and Placement 3 on the 1982 California Achievement Tests--Total Reading	79
19.	Least Significant Difference Procedure for Groups Within Grade 4 and Placement 1 on the 1982 California Achievement Tests--Total Reading	80
20.	Summary of Group Within Grade 3 and Placement 1 on the 1982 California Achievement Tests--Total Mathematics	81
21.	Summary of Group Within Grade 3 and Placement 2 on the 1982 California Achievement Tests--Total Mathematics	81

Table

Page

22.	Summary of Group Within Grade 3 and Placement 3 on the 1982 California Achievement Tests--Total Mathematics	81
23.	Summary of Group Within Grade 4 and Placement 1 on the 1982 California Achievement Tests--Total Mathematics	82
24.	Summary of Group Within Grade 4 and Placement 2 on the 1982 California Achievement Tests--Total Mathematics	82
25.	Summary of Group Within Grade 4 and Placement 3 on the 1982 California Achievement Tests--Total Mathematics	82
26.	Summary of Group Within Grade 5 and Placement 1 on the 1982 California Achievement Tests--Total Mathematics	83
27.	Summary of Group Within Grade 5 and Placement 2 on the 1982 California Achievement Tests--Total Mathematics	83
28.	Summary of Group Within Grade 5 and Placement 3 on the 1982 California Achievement Tests--Total Mathematics	83
29.	Summary of Group Within Grade 6 and Placement 1 on the 1982 California Achievement Tests--Total Mathematics	84
30.	Summary of Group Within Grade 6 and Placement 2 on the 1982 California Achievement Tests--Total Mathematics	84
31.	Summary of Group Within Grade 6 and Placement 3 on the 1982 California Achievement Tests--Total Mathematics	84
32.	Least Significant Difference Procedure for Groups Within Grade 3 and Placement 1 on the 1982 California Achievement Tests--Total Mathematics	86

Table

: Page

33. Least Significant Difference Procedure for Groups Within Grade 3 and Placement 3 on the 1982 California Achievement Tests-- Total Mathematics	86
34. Least Significant Difference Procedure for Groups Within Grade 5 and Placement 3 on the 1982 California Achievement Tests-- Total Mathematics	86
35. Least Significant Difference Procedure for Groups Within Grade 6 and Placement 2 on the 1982 California Achievement Tests-- Total Mathematics	87
36. Least Significant Difference Procedure for Groups Within Grade 7 and Placement 3 on the 1982 California Achievement Tests-- Total Mathematics	87
37. Summary of Group Within Placement and Grade on the 1982 California Achievement Tests-- Total Reading	89
38. Summary of Group Within Placement and Grade on the 1982 California Achievement Tests-- Total Mathematics	90
39. Summary of Level Within Grade 3 and Placement 1 on the 1982 California Achievement Tests--Total Reading	92
40. Summary of Level Within Grade 3 and Placement 2 on the 1982 California Achievement Tests--Total Reading	92
41. Summary of Level Within Grade 4 and Placement 3 on the 1982 California Achievement Tests--Total Reading	92
42. Summary of Level Within Grade 5 and Placement 2 on the 1982 California Achievement Tests--Total Reading	93

Table	Page
43. Summary of Level Within Grade 5 and Placement 3 on the 1982 California Achievement Tests--Total Reading	93
44. Summary of Level Within Grade 3 and Placement 1 on the 1982 California Achievement Tests--Total Mathematics	94
45. Summary of Level Within Grade 3 and Placement 2 on the 1982 California Achievement Tests--Total Mathematics	94
46. Summary of Level Within Grade 4 and Placement 3 on the 1982 California Achievement Tests--Total Mathematics	94
47. Summary of Level Within Grade 5 and Placement 2 on the 1982 California Achievement Tests--Total Mathematics	95
48. Summary of Level Within Grade 5 and Placement 3 on the 1982 California Achievement Tests--Total Mathematics	95
49. Summary of Level on the 1982 California Achievement Tests--Total Reading	97
50. Summary of Level on the 1982 California Achievement Tests--Total Mathematics	98

CHAPTER ONE

Introduction

Discussion

The Department of Research and Evaluation was asked by the Superintendency and, in particular, the Assistant Superintendent for Educational Services at the Elementary Level to investigate the impact of multigrade classrooms on student achievement. The Mesa Public Schools has a history of using multigrade/multiage classrooms as a means of delivering educational goals and objectives to students. This history can be traced back to the original one room schoolhouse in the city of Mesa. Over the years the schools in Mesa followed the national trend to having one grade serve students at a given age. In the 1960s and early 1970s, Mesa, based on research findings and philosophical considerations, again experimented with multigrade/multiage placement in a number of elementary schools in the district. Although these experiments met with reasonable success, multigrade/multiage grouping was not adopted as a district philosophy. At the same time these studies were being conducted the press of finances and class load dictated the use of multigrade classes. The Mesa Public Schools, as many school districts throughout the nation, were forced into using multigrade classes more and more in an attempt to both economize and equalize class loads.

These considerations and evolving utilization of multigrade classes formed the basis for the Assistant Superintendent for Educational Services at the Elementary Level to request a study to be done by the Department of Research and Evaluation. The major question being asked was whether or not placement into

multigrade classes had any detrimental effect on student achievement. In pursuing the answer to this question it was felt that a thorough research study should be conducted; that is, a comprehensive literature review should accompany the investigation of student performance. It is felt that this report does just that: it merges a comprehensive review of the research literature with an indepth study of the academic performance of Mesa elementary students in multigrade classes.

Background

A hundred years ago multigrade classrooms were common features, particularly in the one-room schoolhouse in rural areas. Economics and student enrollment dictated that one teacher would teach all students in the community in all grades. At the time, little attention was paid to whether this grouping method was the best for students; it was the only grouping available.

In the early 1800s, educational, social, political, and economic conditions influenced the development of the graded structure in the elementary school. The establishment of teacher training schools, the popularity of the monitorial system, the influence of German education, the call for state-supported education, and the appearance of graded texts all served to solidify the graded structure for American schools (Goodlad and Anderson, 1963).

No sooner had the graded structure become firmly established in the public school system than critics began to attack its rigidity and lack of individualization. From

the mid-1800s to the mid-1900s, numerous grouping and grading plans were developed to improve instruction and enhance student learning (Shane, 1960).

The philosophy of grouping children into grades was based upon the assumption that greater learning takes place when children of similar achievement are in the same classes. Rehwoldt (1957) suggested that although educational research has not substantiated this practice, it continues because of administrative expediency and the contention that teaching is easier when limited to a single grade.

As a response to the inadequacies of the single-grade organization, especially in an attempt to meet individual needs, nongraded and multigraded organizations were implemented, but their use was limited. Conclusive findings regarding achievement in nongraded and multigraded classes remain to be obtained (Barr and Dreeben, 1977; Lincoln, 1981).

During this same period, from the 1920s to the mid-1930s, ability grouping was used extensively as a method for meeting individual student needs. The National Education Association (1968) reported that although the research on ability grouping was extensive, the issue of its effectiveness was not settled; research results were inconclusive.

The grouping question becomes even more important when coupled with the economic climate in most school districts today. With few corners left to be cut in the budgeting process, hiring one teacher to teach the overflow of students in two grades provides a savings in teacher contracts. But, does achievement suffer? Saving dollars is not justified if student learning is sacrificed. Research is needed to determine whether or not multigrade classes can be justified as a viable alternative to single-grade classes. Multigrade classes are not extolled as the solution to all grouping problems; on the contrary, they are offered as a cost-saving alternative that is as effective for student learning as the standard single-grade class.

Statement of the Problem

The problem in this study was to determine the effects of multigrade classes on student achievement in reading and mathematics. In addition, the achievement of students in differing ability groups was analyzed. The study analyzed the achievement differences in the following areas.

1. Grouping patterns: multigrade classes, single-grade classes in multigrade schools, and single-grade classes not in multigrade schools;

2. Placement: multigrade classes with high-achieving students, multigrade classes with average/high-achieving students, and multigrade classes with average-achieving students;
3. Levels: upper--a fourth-grade student in a third/fourth multigrade class--or lower--a fourth-grade student in a fourth/fifth multigrade class; and
4. Grades: third, fourth, fifth, and sixth.

The following are the specific research questions to be answered:

1. Is there a significant difference among the sample means of three grouping patterns as measured by reading and mathematics achievement tests?
2. Is there a significant interaction among the sample means of three grouping patterns and four different grades as measured by reading and mathematics achievement tests?
3. Is there a significant interaction among the sample means of three groups and three placement types as measured by reading and mathematics achievement tests?
4. Is there a significant difference among the sample means of two levels as measured by reading and mathematics achievement tests?
5. Is there a significant interaction among the sample means of three placement types, two levels, and three grades as measured by reading and mathematics achievement tests?

Definition of Terms

Ability grouping: The grouping in which students of given grades or levels are assigned to classes based upon intelligence, achievement, teacher judgments, or a combination of any of these criteria.

Group 1: The students included in the study from multigrade classes.

Group 2: The students included in the study from single-grade classes in schools having multigrade classes.

Group 3: The students included in the study from single-grade classes in schools not having multigrade classes.

Heterogeneous grouping: The grouping in which students of a given grade or level are assigned to classes without using achievement, intelligence, or other criteria.

Homogeneous grouping: The ability grouping in which students are assigned to classes based upon intelligence, achievement, or other criteria.

Horizontal school organization: The school organization involved with placement of students in instructional groups, such as ability grouping, departmentalization, and team-teaching.

Level: The upper or lower level of students in a multigrade class. The third-grade students in a second/third multigrade class are in the upper level, and the

third-grade students in a third/fourth multigrade class are in the lower level.

Mathematics achievement: Mathematics achievement is the normal curve equivalency score on the California Achievement Tests, Total Mathematics.

Multigrade class: The classroom organization that groups students from two grades in which students retain a specific grade assignment.

Normal curve equivalency (NCE): An equal-interval standard score ranging from 1 to 99, with a mean of 50 and a standard deviation of 21.06.

Placement: The ability grouping used to place students in multigrade classes--high placement, multigrade classes with high-achieving students; average/high placement, multigrade classes with average/high-achieving students; and average placement, multigrade classes with average-achieving students.

Reading achievement: Reading achievement is the student normal curve equivalency score on the California Achievement Tests, Total Reading.

Self-contained classroom. The most common basis of organization in elementary schools--indicates a single teacher assigned to a classroom and responsible for most instruction.

Single-grade class: The classroom organization that groups students of the same chronological age at only one grade or level for instruction.

Vertical school organization: The school organization concerned with the advancement of students through the curriculum, grades, or levels of a school's program, such as single-grade, non-graded, or multigrade programs.

Limitations

1. Though this study has an experimental design, it is not a true experiment because the "sampling by convenience" method is used. To minimize this disadvantage, a demographic analysis will be performed on student characteristics to ensure comparable samples.
2. Differences in the competency of individual teachers are recognized but will not be investigated.
3. Student and parent attitudes toward student placement in a multigrade classroom may affect achievement, but they will not be surveyed in this study.
4. The instructional techniques of individual teachers will not be investigated.
5. Generalizations to other populations must consider the community and school district descriptions cited in Chapter Three.

Purpose of the Research

Rosenbaum (1980) suggested that the issue of student diversity continues to challenge educators. One of the main responses to student diversity has been to divide students into more homogeneous groups--first into "grades" and later into ability and curriculum groupings. Rosenbaum suggested that this grouping may better meet student needs, but there remains the concern for social segregation and social inequality. Thus, educators are caught in a dilemma.

A major deficiency of research on grading plans has existed because of the confusion between vertical and horizontal school organization. Goodlad and Anderson (1963:214) attempted to clarify the misunderstanding. In doing so they stated:

Vertical school organization is designed to serve the necessary school function of moving pupils upward from the time they enter the school unit until the time they leave. There are two polar, alternative structures for doing this: the graded and the nongraded systems. Between these two systems lies a modification of grading known as multi-grading.

In addition to deciding on a vertical plan, school administrators must decide upon a horizontal plan, the method of grouping students within the vertical plan. Students may be grouped horizontally according to homogeneity in ability, achievement, or interests, or they may be grouped heterogeneously (Franklin, 1967a).

A deficiency in the research has been to investigate grading plans but to ignore the method of horizontal, or interclass, grouping. Consequently, there has been no way of knowing whether or not significant differences were related to the vertical or to the horizontal school organization. The purpose of this study was to investigate both types of organization. The first part of the study analyzed the differences between single-grade and multigrade classes (vertical organization). Once a difference was either accepted or rejected, further analysis determined the effects of different methods of placement (horizontal organization).

Differences among grades were included in some multigrade research (Rehwoldt and Hamilton, 1957; Chace, 1961; Finley and Thompson, 1963; Yerry, 1964). Level is the upper or lower grade in a multigrade class; for example, third-graders in a third/fourth multigrade class are in the lower level. This additional information on grade and level can help school administrators plan effectively for using multigrade classes.

Because research on the cognitive effects of multigrade classes has been sparse and inconclusive (Chace, 1961; Bidwell and Kasarda, 1975; Hollifield, 1980) and the research on ability grouping has been extensive and inconclusive (Otto, 1964; Adair, 1978; Barr and Dreeben, 1977;

Lincoln, 1981; National Education Association, 1968; Borg, 1966; Yates, 1966; Findley and Bryan, 1975), the issue of gradedness and ability grouping deserves further attention. Despite this nebulous situation, educators continue to group and, in the practical world of the elementary principal, judgments regarding grouping are based more upon individual bias than upon research-supported results.

The issue may have come full circle. In the one-room schoolhouse, grouping was a financial necessity, and, given the budgetary restrictions existing today, that aspect of the one-room schoolhouse--combining grades--may be as appropriate today as it was in the 1800s. For economic reasons, in both increasing and declining enrollment districts, many districts are again using multigrade classrooms, usually combining only two grade levels, e.g., four and five, five and six (Lincoln, 1981). Even the popular press ("But They Can't Close MY Child's School!" 1980) has suggested multigrade classes as a solution for economic difficulties.

Within these multigrade classrooms, different approaches to ability grouping are used. It has been suggested that multigrade classes with differing ability grouping may achieve the same academic achievement as the more common single-grade heterogeneous classes. Goodlad and Anderson (1963) suggested that because vertical school

organization is merely a device used to facilitate placement of subject matter, student progression, and the setting of standards, there is little chance that organization design alone will change student outcomes.

The results of this study will be helpful to classroom teachers, administrators, parents, and other persons interested in using multigrade classes for economic reasons, which were cited in the 1960s research on vertical organizational structure (Anderson, 1962). Research is needed to assure the decision makers in education that there are no detrimental effects on the academic growth of students placed in multigrade classes.

Summary

This chapter outlines an investigation of the effects of multigrade grouping on the reading and mathematics achievement of elementary students. It presents background for the study, the statement of the problem, definition of terms, limitations, and purpose of the research.

Chapter Two includes a review of related literature and research. Chapter Three presents instruments and design of the study. Chapter Four presents the results and findings of the study, and Chapter Five includes the summary, conclusions, and recommendations.

CHAPTER TWO

Review of the Research

An interest in grouping was focused upon quite early by educators. The graded structure of the mid-1800s developed out of a need to classify the many children who were beginning to attend school as a result of increased interest in public schooling. Years later, in the Thirty-fifth Yearbook of the National Society for the Study of Education, Whipple (1936:1) presented a statement of the importance of grouping that holds true today. He stated the educational belief

that children do present problems of individual differences; that it is necessary to make provision for them; that a suitable educational philosophy must be developed; that general adaptations of curriculum and teaching method at the various levels of learning ability must find their way into educational processes; and that the class groups in which children are placed have an important bearing on these problems.

In order to place the current state of school organization into proper perspective, a historical overview of grading plans and grouping procedures since the early 1800s is presented. The review of research is divided into three sections: research on nongraded versus single-graded plans, research on multigraded versus single-graded plans (methods

of vertical organization), and research on ability grouping procedures (a method of horizontal organization).

History of Grouping/Grading

Individualized instruction was the procedure used in the "district" schools of the eighteenth century and the dame schools of the seventeenth. Though teachers were poorly prepared and the curriculum was haphazard, the instruction was highly individualized (Beggs and Buffie, 1967; Goodlad and Anderson, 1963; Morgenstern, 1966). In the district school, pupils of all ages met in one room with one teacher for instruction; hence the term one-room schoolhouse emerged. Because of sparse population and limited finances, the district school continued to dominate education until changed by rapid urban growth at the beginning of the nineteenth century.

Other changes were occurring that would influence grouping. The first normal schools were established in the United States in 1823 and 1827, and these early teacher-training institutions encouraged the belief in the graded school system. In Boston in the eighteenth century, separate reading and writing schools were developed; and a defining of curriculum began to occur. Specific material was to be learned by a certain age; thus, "norms" were being developed (Beggs and Buffie, 1967).

The monitorial, or Lancastrian, system brought to New York in 1806 made teaching large groups of children more manageable. A teacher would teach a lesson to older students, or monitors, who in turn would impart this knowledge to a smaller group of students. The system necessitated a careful process of grading so that students at the same level could be placed together (Morgenstern, 1966; Otto and Sanders, 1964). The low cost of the monitorial system and its emphasis upon group instruction blended well with the broader educational, social, political, and economic conditions of the time described by Goodlad and Anderson (1963:46):

Meanwhile, growing confidence in the capacity of the human race for unending progress, the diffusion of religious humanitarianism, the beginning of the labor movement, and growing nationalism created a milieu that was receptive to the revolutionary idea of education for all. In such a setting, the ideas of educational spokesmen like Mann, Stowe, and Barnard, while controversial, struck many receptive ears. Each section of the country had its nineteenth-century leaders who attacked the dogma that education is the privilege of the few who can afford tutor, grammar school, or academy and not the business of government. The graded system was to provide an orderly means of classifying the many young people who were to come to school as a result of increased public interest in schooling.

In 1848 the Quincy Grammar School of Boston offered a more defined pattern of graded class structure than ever before seen in education. The principal of the school reorganized the school based upon the German model described by Horace Mann in 1844. One important element of this

reorganization was to categorize by chronological age the large number of students now enrolled in public school (Otto, 1934).

Otto and Sanders (1964) stressed that the graded school was more than an effort to arrange pupils. The grading included the designation of minimum essentials or standards for each grade. Thus, the graded school maintained the philosophy that specific subject matter and skills could be developed and graded in a set sequence.

Another factor contributing to gradedness was the appearance of graded textbooks. The success of the McGuffey Eclectic Readers in 1836, arithmetic texts in 1821, and later texts in spelling, science, social studies, and grammar solidified the importance of graded materials (Beggs and Buffie, 1967; Goodlad and Anderson, 1963; Otto and Sanders, 1964).

Goodlad and Anderson (1963:48) cited five major developments that affected the graded structure in American schools:

- (1) The movement toward public state-supported education;
- (2) the practical success and astonishing economy of the monitorial system;
- (3) the several appeals of German education as interpreted by American spokesmen;
- (4) the call for trained teachers, answered in part by the founding of state-supported normal schools;
- and (5) the appearance of graded school books in all areas of instruction.

No sooner had the graded structure become established in the cities than educators began to question its

effectiveness. Critics contended that graded classes were too rigid and that individual differences were not being met, as evidenced by the large numbers of students who were not being promoted. Shearer (1899) pointed out the absurdity of expecting children to be at the same stage of mental development in a given grade. He offered a Procrustean analogy: If an average height were to be obtained for all fourth-grade children, then those who were naturally tall would be compressed and short ones would be stretched. He suggested that the schools had limited mental development in the same way by gearing curriculum to the "average pupil."

Because of this concern and others like it, several school districts created plans to better meet the student needs and avoid nonpromotion. In 1862, W. T. Harris developed the St. Louis Plan as a modification of the graded system. Though still a graded structure, it included more frequent promotion and reclassification. These were at ten-week intervals instead of the accepted annual promotion (Goodlad and Anderson, 1963; Beggs and Buffie, 1967). Preston Search initiated the Pueblo Plan in 1888 to encourage individual progress for each pupil. Course content was organized for pupils to complete different amounts of work at different rates. In addition, no marks were given and

nonpromotion was eliminated (Otto, 1934; Beggs and Buffie, 1967).

The Cambridge Plan and the Portland Plan attempted to provide differing rates of learning for groups of children. The Cambridge Plan was developed as a nine-year elementary school. All students would complete the same work during the first three years, but students could complete grades four through nine in four, five, six, seven, or eight years. The plan was intended primarily for average and superior students (Otto, 1934; Goodlad, 1963; Beggs, 1967). In the Portland Plan, developed in 1897, nine grades were divided into fifty-four units. Average students covered six units in a year, and bright students covered eight per year. The units were arranged so students could be transferred from the regular- to the rapid-progress group, and vice versa, as needed (Otto, 1934).

In Batavia, New York, in 1898, John P. Kennedy introduced a plan to employ additional teachers to give special assistance to slow learners. This was one of the first plans implemented to provide help for low-achieving students (Beggs and Buffie, 1967; Morgenstern, 1966). The North Denver plan was the reverse of the Batavia plan, with bright students receiving special help. All students covered a core curriculum, but bright students received enrichment activities (Otto, 1934).

Both the Dalton Plan and the Winnetka Plan emphasized an individualized task approach to curriculum organization. In 1919, Carleton Washburne developed the Winnetka Plan. Washburne's course of study was divided into basic subjects that everyone was expected to master, and into group and creative activities developed on an individual level. Closely related was the Dalton Plan, introduced in 1919 by Helen Parkhurst, which emphasized the individual's performance in group situations. Subjects were divided into academic and nonacademic groups and were organized sequentially so that students progressed on an individual basis. The work for each grade was described in contracts that defined the content and time limits (Beggs and Buffie, 1967; Otto and Sanders, 1964; Goodlad and Anderson, 1963).

The 1930s held the beginnings of the nongraded movement. Goodlad and Anderson (1963) defined the nongraded school as one in which grades and grade labels are replaced by a plan of continuous progress. It is intended to eliminate the promotion-nonpromotion issue of graded schools with the school organization adjusting to student needs rather than the student conforming to school structure. The Flexible Progress Plan in 1934 typified this philosophy. Grades one, two, and three were abolished to allow for greater individualization. Milwaukee was the locale of the first large school district to use the nongraded plan, which

was begun in 1942 by Lowell Goodrich and represented a major step in the nongraded school movement (Beggs and Buffie, 1967). The Joplin Plan in 1953 instituted a nongraded reading program. At the given reading time, students in grades four through six would break into different groups for reading instruction, followed by twenty-five minutes of free reading (Brown, 1965).

Except for a few innovative attempts, the nongraded school spread very little until the mid-1950s. In survey data collected by Goodlad and Anderson (1963) in 1958, fifty nongraded schools were documented. Two years later, eighty-nine communities had nongraded schools. A National Education Association Project on Instruction survey reported that 26 percent of the principals expected to have some form of nongradedness by 1966 (Miller, 1967). Though this interest was not high, it did show a growing interest. The high interest in nongradedness of the sixties waned during the seventies.

The multigrade, or multiage, classroom was a common feature in the one-room schools in the 1800s; in fact, it was not challenged as the organizational arrangement of the elementary school until the graded school was established in 1848. By the 1860s the graded school had replaced the multigraded school except in rural areas. After many attempts to modify the graded school while maintaining

the grade structure, multigrade programs emerged (Lincoln, 1981). The multigrading of the 1930s had its origin in the Jenaplan introduced in 1924 in the Jena University Elementary School in Germany (Dawson, 1961).

From the 1930s through the 1970s multigrading was espoused as a clear improvement over single-grading. Chace (1961) cited philosophical, developmental, psychological, educational, and administrative criteria to support his case for multigrade grouping. Franklin (1967b:69) described the philosophy of multigrading as proposed in the 1960s:

The multigrade plan is based on the assumption that students are better off when they are placed in a class on the basis of their differences rather than their likenesses with sub-grouping within the class. Multigrading operates within a framework of a graded system, and grade labels are retained. The plan is comparable to the one-room school which preceded the graded school of 1848. It is one alternative to the graded system. However, it must be recognized that multigrading has not had wide acceptance.

Multigrading has the advantage of providing an organization that tends to emphasize individual differences while the use of grade labels provides a nomenclature familiar to parents. The multigrade plan has been used in diverse places: Torrance, California (Hull, 1958), Chicago, Illinois (Gilbert, 1964), Englewood, Florida (Goodlad and Anderson, 1963), Minneapolis, Minnesota (Franklin, 1967), and East Brunswick, New Jersey (Anderson, 1962).

Thus, the attempted changes in school organization can be seen to have been numerous. Shane (1960) cited

thirty-two grouping plans that have been used in education, and the National Education Association (1968) presented twenty-six plans. Over the years many of these plans have been discarded, modified, or gradually accepted. The progression of organizational patterns has moved from the ungraded one-room schoolhouse to well-defined grades with set curriculum. Though many variations to graded classes have surfaced since the mid-1800s, the graded class with some type of intraclass grouping has prevailed.

Research on Nongrading Versus Single-Grading

In the Encyclopedia of Educational Research in 1970 and again in 1982, the authors concluded that the research favors nongrading. Heathers (1969) commented on the paucity of research studies on nongrading. In fact, most research was conducted in the 1960s and early 1970s, with little research after 1973 (Heddens, 1982). A review of eight studies on nongrading for the time period of 1959 to 1965 was presented by DiLorenzo and Salter (1965). Six studies made comparisons in reading achievement in grades three through six. Four of these studies found the achievement of nongraded students significantly higher than that of graded students, one study noted no difference, and one found the graded students to be significantly higher in achievement than the nongraded. Of five studies that made comparisons

in mathematics achievement, four noted significantly higher results for the nongraded students; the remaining study favored graded classes.

From a review of the research between 1968 and 1971 comparing graded and nongraded elementary schools, Pavan (1973) concluded that achievement test scores favored nongradedness. Pavan cited sixteen studies that used standardized achievement tests to compare graded and nongraded schools. Of these, only one comparison favored the graded school. All others favored the nongraded school or indicated no statistically significant differences between groups. Based upon these data, Pavan stated that there should no longer be concern that placing children in nongraded programs will be detrimental to their academic achievement. In a later review of research on nongraded programs, Martin and Pavan (1976) cited nine studies showing achievement results favorable to nongraded programs. The reviewers concluded that when students were matched for intelligence, the nongraded achievement scores were significantly higher.

A check of some of these reviewed studies indicated that their findings were loosely reported. Many of the nongraded classes had higher scores than single grades, but some were not statistically significant. Reviewers tended to pick out the results of the research that supported their

claims. For example, in a study by Bowman (1971), test scores of intermediate students favored nongraded organization, but scores for primary students did not. In this case the Bowman study was said to favor the nongraded organization, which is partially true. Though the literature on nongraded programs does favor the nongraded program, the results are not conclusive. Studies by Morris (1969), Hillson et al. (1964), Halliwell (1963), Ward (1970), and Killough (1971) favored the nongraded program; a study by Carbone (1961) favored the graded program; and studies by Remacle (1971) and Bowman (1971) had mixed results. These studies are explained further in the following text.

Few longitudinal studies on nongraded programs have been conducted; the three cited (Morris, 1969; Ward, 1970; Killough, 1971) favored the nongraded organizations. One longitudinal (five years, including a postexperimental phase) study (Morris, 1969) was conducted to determine the achievement differences for students in graded and nongraded classes. The 60 sixth-grade students had been in a graded structure in the primary grades; the 57 fifth-grade students had experienced nongraded classes in grades one through three (Stage I). In grades four and five (Stage II), the grouping was similar for both groups. An analysis of covariance was performed for both stages partitioned by method and sex, with intelligence as the covariate. The

mean I.Q. for the graded group was 116.8; for the nongraded group, I.Q. was 114.5. The analysis of covariance adjusted achievement scores for Stage I and Stage II were higher for nongraded students. Thus, the experimental gains in Stage I were maintained during Stage II. Also, although analysis of covariance could not adjust for the difference in I.Q., the results were stronger because the lower-I.Q. group (nongraded) had the higher achievement.

Another longitudinal study on nongraded school organization was presented by Killough (1971). Over a three-year period pretest and posttest achievement data were collected on three hundred elementary pupils who were equally divided by sex and year in school. One half of the students were enrolled in a nongraded open-space facility, and the other half were in a single-grade traditional facility. A multivariate analysis of variance design was used to analyze the independent and interactive effects that sex and type of school program had on the cognitive achievement gains of the subjects in arithmetic reasoning, concepts, and computation, and in reading comprehension and vocabulary. The findings suggested that students who remained in a nongraded program for three years had significantly higher mean achievement gains in most cognitive areas than students in a single-grade program; girls in a nongraded program achieved at a faster rate in arithmetic computation than boys; boys in

a nongraded program achieved at a faster rate in reading comprehension than girls.

A third longitudinal nongraded study by Ward (1970) included first-grade students the first year and second-grade students the second year. A total of 797 students and 27 teachers were involved in the program over the two-year period. Standardized tests for achievement in reading, arithmetic, and readiness were analyzed with the t-test. The nongraded program was favored in all findings.

Two other studies (Halliwell, 1963; Hillson et al., 1964) favored the nongraded organization. Halliwell (1963) investigated the effects of placement in nongraded primary classes on achievement. He compared the achievement of primary students who had been taught for one year in a nongraded program in reading and spelling with the achievement of students taught in a graded class. No attempt was made to adjust the curriculum in the nongraded program; only the organizational structure was changed. Intelligence scores were not significantly different between the two groups at any grade level. The simple t-test was used to compare spring achievement test results of the two groups. The grade-equivalent scores of the nongraded students were higher in thirteen of fourteen subtests for the three grade levels. The differences were found to be statistically significant for six of the subtests. Interestingly, half of

the significant differences were for mathematics, which was not taught in a nongraded organization. Halliwell did state that teachers spent less time on reading instruction in the nongraded group and more time on arithmetic, social studies, and language arts.

Hillson et al. (1964) conducted a study of first-grade students placed in both graded and nongraded classes. Hillson has been credited by others for taking care to design an experiment that would control the variables of teachers, instructional materials, and in-service training. Achievement results were analyzed at the end of one and one-half years using the t-test. The students in the nongraded organization scored significantly higher on two reading achievement measures.

The only study that supported graded over nongraded classes was conducted by Carbone (1961). He compared the achievement of 122 intermediate students who were taught in a nongraded primary program, with 122 intermediate students who were taught in a graded primary program. Four school systems that were similar in population, socioeconomic level, and geographic location were included in the study. Students were matched by age and sex, and I.Q. was held constant by using analysis of covariance. Achievement in reading and language was analyzed using the t-test. In all achievement areas the student scores from the graded classes

were found to be significantly higher than those from the nongraded classes. Goodlad and Anderson (1963) criticized the Carbone study because no significant differences were reported in instructional practices between the graded and nongraded programs. Goodlad and Anderson assumed that nongrading produces changes in instructional practices; if none occur, there is no reason to expect any change in instructional outcomes.

Two additional studies of nongraded programs (Remacle, 1971; Bowman, 1971) resulted in mixed results. The Remacle research studied the effects of nongraded and graded organization on fifth- and sixth-grade students. Remacle used the Iowa Test of Basic Skills as the achievement measure for sixty-four nongraded and sixty-four graded students; the statistical test used was not disclosed. In the achievement areas of vocabulary, reading, language, arithmetic, and composition, the sixth-grade nongraded students consistently scored higher than the sixth-grade graded students. Only the area of work study-skills resulted in a statistically significant difference, and this difference favored the nongraded group. No significant differences were found in the achievement of fifth-grade students, although scores were higher for the nongraded group.

Bowman (1971) investigated the differences in achievement of graded and nongraded students using the Metropolitan

Achievement Test as the measurement variable. Analysis of covariance was used to test for difference, with intelligence used as the covariate. No significant differences were found between the nongraded and graded organizations for the primary level, but at the intermediate level students in the nongraded organization scored significantly higher. The author suggested that though nongrading attempts to provide a favorable setting and conducive climate for some problems in elementary schools, it in no way guarantees improvement for all elementary schools.

Though these research studies on the nongraded organization tended to be favorable, the conclusion of DiLorenzo and Salter (1965) applies: the research does not warrant a decision to adopt a nongraded organization primarily because of the short experience with nongrading. Also, as Wilt (1971) concluded, "no single organizational change can hope to solve the needs of learners in a mass, heterogeneous society."

Research on Multigrading Versus Single-Grading

The limited research on multigrade classes began in the 1930s and has continued to the present. Only fourteen studies were located that investigated achievement effects of elementary multigrade classes. Most of the research studies found no significant differences between multigrade

and single-grade achievement scores. All of the studies used standardized tests as their achievement measures; eleven of the studies used reading and mathematics scores. Seven studies limited the grade span to primary (Harvey, 1974; MacDonald and Wurster, 1974; Adair, 1978; Lincoln, 1981) or intermediate (Foshay, 1948; Drier, 1949; Adams, 1953) grades; seven studies included both primary and intermediate grades (Knight, 1938; Rehwoldt and Hamilton, 1957; Chace, 1961; Finley and Thompson, 1963; Yerry, 1964; Way, 1969; Milburn, 1981).

Only two studies (Yerry, 1964; Milburn, 1981) considered the achievement of lower- and upper-level students; in other words, these studies sought to determine if second-graders achieve differently in a first/second multigrade class than they do in a second/third multigrade class. None of the studies considered differing ability for placing students in the multigrade classes.

Nine studies (Knight, 1938; Drier, 1949; Adams, 1953; Chace, 1961; Way, 1969; Harvey, 1974; MacDonald and Wurster, 1974; Adair, 1978; Lincoln, 1981) showed no significant difference in the achievement of multigrade and single-grade students. The earliest investigation of multigrade classes was conducted by Knight (1938). The multigrade classes were evaluated with regard to student educational achievement and problems concerning the organization, administration, and

teaching of double grades. Seven third/fourth multigrade classes were matched with six single fourth-grade classes; seven fourth/fifth multigrade classes were matched with the same six single fourth-grade classes. The findings indicated that on the complete Stanford Achievement Test fourth-grade children in multigrade classes equaled or surpassed in gains the children in rooms with only fourth-grade pupils, but not at a statistically significant level. Conclusions regarding other than achievement data were as follows: (1) multigrade classes were formed as an economy measure with no standard for assigning pupils; (2) principals and teachers were not in favor of multigrade classes, while children did not have a preference; (3) except for reading and mathematics, teachers taught their classes as if they had one grade; and (4) the teacher, rather than the grouping, appeared to be the important factor.

A comprehensive and well-designed study on single-grade and multigrade classes was conducted by Drier (1949). The purpose of the study was to determine whether or not differences existed between the achievement of rural children from graded and ungraded elementary schools as measured by reading, language, arithmetic, and spelling tests at the sixth-grade level. Forty-one counties were selected for participation in the study. Because of the large number of ungraded schools that had sixth grades of only a single

student each, the county was the unit for all comparisons. A total of 599 graded and 923 ungraded sixth-grade students participated in the study. The t-test was used to determine the significance of differences in achievement. Sixth-grade students did not differ significantly ($p < .05$) on reading or mathematics.

Another study on multigrade classes finding no significant differences considered student achievement in nineteen multigrade and nineteen single-grade classes. Adams (1953) matched 150 fifth-grade students who had been placed in a fourth/fifth multigrade class with 150 students in regular fifth-grade classes. Mean gains in reading, arithmetic, and language on the Progressive Achievement Test were computed, but no differences were statistically significant between the combination classes and the single-grade classes. The statistical procedures were not reported.

Chace (1961) investigated multigrade units in which two to four different grade levels were represented and taught by one teacher. The study centered around academic development of students, and parent, teacher, and administrator perceptions of multigrade groups. Students were matched by sex, age, I.Q., grade placement, socioeconomic characteristics of the schools, and professional training and experience of the teachers. The classes of three teachers in a campus laboratory school were matched with classes of

fifty-seven teachers in public schools. Conclusions made from the Stanford Achievement Test results were that multigrade grouping affected achievement positively, but not statistically significantly, at grades three to six; attitudinal data revealed that parents accepted the theory of multigrade grouping, but not its practice; and that the success of multigrade grouping was limited by the continued use of traditional graded activities. Chace stated that the chief implication for multigrade groupings is that the barriers of single-grade grouping must be removed for multigrade grouping to have differing results. It should be noted that because Chace's study compared students in public schools with those in a laboratory school, generalizations were limited.

Way (1969) completed a study on multigrade students in a suburban New York school district. A total of 671 children in single-grade classes in grades one through five formed the control group; 135 students from combination classes completed the experimental group. The multigrade teachers had taught in multigrade classrooms for two or more years. The Comprehensive Test of Basic Skills was the achievement measure for the study. A 2x2 analysis of variance factorial design was used to analyze the data. The two experimental variables were type of classroom and sex; the dependent variable was the achievement score. No

significant differences were found on the effects of classroom type, sex, or interactions of classroom type and sex.

The purpose of a study by Harvey (1974) was to analyze the effects of multigrade grouping on readiness development and achievement at the kindergarten level. The experimental group involved 31 students in three multigrade kindergarten/first-grade classes and the control group consisted of 152 kindergarten students in seven single-grade classes. The achievement measures were the Metropolitan Readiness Test and the Stanford Early School Achievement Test. Based upon comparisons between population means, the author concluded that the classroom organization (multigrade or single-grade) had little or no effect on readiness or achievement.

Another study including kindergarten/first-grade classes used first grade as the experimental unit. Achievement comparisons were made between kindergarten and first-grade students in single-grade and multigrade classes in Adair's (1978) ex post facto comparative design. Age, sex, mental ability, socioeconomic status, and teacher education level were used as covariates. Achievement of first-grade students was measured using the Cooperative Primary Tests; no cognitive measure was used with kindergarten. Data were analyzed using an analysis of covariance technique. The sample consisted of 453 kindergarten and 500 first-grade students. No significant differences were found for

listening, word analysis, reading, or mathematics skills between first-graders in single-grade classes and those in multigrade classes. Adair qualified the achievement results by stating that the achievement tests used could have been too general to measure achievement differences.

MacDonald and Wurster (1974) conducted a study on the achievement effects in reading on first-grade students when a multigrade primary team (first, second, third combined) was replaced by a single first grade. The study included data comparing second-graders who spent first grade in a multigrade team with second-graders who were in a segregated first grade. The means of the raw scores of the Gates-MacGinitie Reading Test were compared and tested for significance by the t-test. Findings showed no significant differences in the means.

The most recent studies on multigrade classes were reported by Lincoln (1981) and Milburn (1981). The Lincoln study sought to discover the effects on reading achievement of multigrade classes. Four hundred two children who completed grade two in 1977, 1978, and 1979 were subjects for the study. The Comprehensive Test of Basic Skills and the Short Form Test of Academic Aptitude were the standardized tests used. Tests for differences among groups were performed using analysis of covariance and t-test procedures. The reading achievement test scores did not vary

significantly for the single-grade and multigrade groups. The author concluded that a multigrade primary program did not hinder reading achievement of children.

Four multigrade studies (Rehwoldt and Hamilton, 1957; Finley and Thompson, 1963; Yerry, 1964; Milburn, 1981) reported results that were mixed. Three of the studies (Rehwoldt and Hamilton, 1957; Yerry, 1964; Milburn, 1981) indicated results that favored multigrade at some grades or on some tests and showed no significant differences on other grades or tests. The fourth study (Finley and Thompson, 1963) reported some results that favored multigrade classes and other results that favored single-grade classes, as well as some results displaying no significant differences.

Rehwoldt and Hamilton (1957) and Hull (1958) reported on a study designed to explore the characteristics of multigrade grouping. The study was limited to four primary multigrade classes and three intermediate multigrade classes; teachers were randomly assigned, and no training was provided. Two hundred ninety matched pairs (age, intelligence, achievement) were used for the comparisons. The types of statistical analyses used in the study were unclear. The reading achievement results favored multigrade programs at first and fifth grades; arithmetic results favored multigrade programs at first, third, and fifth

grades; and language results showed no significant differences.

Finley and Thompson (1963) made a comparison of the achievement of multigrade and single-grade rural school children. The authors hypothesized that there would be no difference in achievement between the two groups, which included fifty-three paired boys and fifty-one paired girls at third grade and sixty-two paired boys and forty-six paired girls at fifth grade who met the criteria for matching. The two groups were matched by sex, I.Q. (California Short Form Test of Mental Maturity), chronological age, and availability of achievement test scores (California Achievement Battery, Form W). Mean differences and t-tests were computed for boys, girls, and total group. At the third-grade level a significant difference was noted for arithmetic fundamentals that favored the multigrade class, but at the fifth grade a statistically significant difference in arithmetic fundamentals favored the single-grade boys. Because the differences were not consistent, generalizations could not be made regarding achievement in multigrade and single-grade classes.

Yerry's (1964) study was important in that it was one of the two studies that investigated not only the difference between multigrade and single-grade classes, but also differences in groupings within the multigrade classes.

Five hundred students in grades one through six from two schools comprised the twenty-two multigrade classes. The control group consisted of 500 children from single-grade classes of similar enrollment, socioeconomic background, and experience of teacher. Pretest and posttest scores on the California Achievement Tests were the achievement variable. A one-way analysis of variance was used to compare the difference between single-grade and multigrade classes. The data analysis showed no consistent pattern of change: (1) no significant differences existed between single-grade and multigrade classes in grades two, three, or six; (2) significant differences in arithmetic ($p < .05$), language ($p < .05$) and total achievement ($p < .01$) for grade one favored multigrade classes; and (3) significant differences in reading ($p < .05$), arithmetic ($p < .01$), and total achievement ($p < .01$) for grade five favored multigrade classes.

The second analysis by Yerry considered the differences between grades within the multigrade classes. Yerry's findings indicated no significant differences in achievement gains for students grouped in first/second, fourth/fifth, or fifth/sixth multigrade classes. In the second/third multigrade class, achievement gains were significantly different for the second- and third-graders. The gains favored the second-grade children in language; the gains

avored the third-grade children in reading, arithmetic, and total achievement. Comparisons of achievement gains of second-grade students in first/second and second/third multigrade classes indicated significantly different achievement gains in all four achievement measures, which favored second-graders in the first/second multigrade classes. No differences were noted for achievement gains by fifth-graders in either fourth/fifth or fifth/sixth multigrade classes.

The Milburn (1981) study included two schools of similar size over a five-year period. The experimental school had five multigrade classes with an average of twenty-five students in each; each class covered a grade span of three years. The two schools were located in similar socioeconomic areas in the suburbs of two cities. Two achievement measures were used: the Gates-MacGinitie Reading Test and the California Achievement Test in Mathematical Computation. Multigrade students scored significantly higher on the vocabulary section of the reading test. The statistical procedures were not included.

The performance of the lower- and upper-level students in each multigrade class was analyzed. In all classes the lower-level students scored higher on the achievement measure than did the students of the same respective grades in single-grade classes. The upper-level students scored

much like their counterparts in the single-grade classes. Milburn concluded that the multigrade grouping promises no significant gains in basic skills; it is simply one of a number of strategies for children. "It is neither the 'ideal' method nor is it a panacea for all educational ills" (Milburn 1981:309)

Only one study (Foshay, 1948) reported significantly different results in favor of the single-grade classes. Foshay evaluated a multigrade class in grades four, five, and six. The school was not typical in that the mean I.Q. of the experimental group was 123 and that of the control group was 129. The control group made significantly greater gains in reading, arithmetic, spelling, and language usage than did the experimental group. The mean gain of the one-year program was approximately one-half year on achievement tests. Generalization was cautioned against due to the small sample.

Ability Grouping

Ability grouping is a form of homogeneously grouping students in a horizontal school organization. Pupils may be grouped on the basis of achievement test results, intelligence test scores, chronological age, interests, behavior, or any combination of these and other criteria. Turney (1936) stated that the aim of ability grouping is to bring

together pupils who will be able to work together and to progress together under conditions permitting the fullest possible development of the individuals involved.

Ability grouping was very prevalent in the United States in the 1920s and mid-1930s. In the 1950s there was renewed interest in ability grouping. The National Education Association (1968), Esposito (1973), and Martin and Pavan (1976) presented the pros and cons of ability grouping. Favorable arguments purported that homogenous ability grouping:

1. Allows pupils to progress at their own rates;
2. Offers the use of methods and materials that are appropriate to the total group;
3. Offers pupils more individual attention from the teacher; and
4. Offers instruction geared to the average of the group rather than to the appropriate individual level.

Critics of homogeneous grouping argued that heterogeneous grouping:

1. Avoids the stigma of placement in a low group;
2. Allows students to work with pupils of differing ability;
3. Permits lower-achieving students to learn from higher-achieving students;

4. Encourages teachers to note student differences and teach toward them;
5. Does not rely on measurement factors that are not accurate enough for placement, particularly based upon only one factor; and
6. Does not segregate children by socioeconomic level or race.

The National Education Association (1968) emphasized the care that must be taken when discussing research related to ability grouping. The variables used in the research often confound any interpretation of the results. Goldberg, Passow, and Justman (1966) listed nine difficulties in generalizing from the research on ability grouping: (1) variation in aim and grouping; (2) differences in number of students, number of groups, and sizes of classes involved; (3) differences in basis for determining homogeneity; (4) differences in duration; (5) differences in the ways of matching experimental and control groups; (6) differences in the treatment; (7) differences in teacher assignments; (8) differences in instruments and techniques used in evaluation; and (9) failure to assess the effects of grouping on teachers and administrators.

Grouping is also intended to reduce the range of variability of achievement in classrooms to simplify instruction for the teacher. Clarke (1958) undertook a study to reduce

the variability in achievement that occurs in ungrouped classes. Students were grouped on the basis of intelligence and on the basis of reading achievement. The data were then analyzed to determine the reduction in variability. The sample included 621 students in five Canadian elementary schools at grade three. The author determined the variability of the ungrouped classrooms, then regrouped the pupils on the basis of ranked I.Q. scores, dividing the students into classrooms of equal size. Variability in I.Q. and achievement were calculated for the new groupings, with the same procedure followed for grouping by reading achievement. Grouping by I.Q. reduced reading and math variability by 9 percent and 2 percent, respectively. Grouping by reading achievement reduced the variability of arithmetic and language scores by 20 percent and I.Q. variability by 10 percent. The results indicated that the reduction in variability achieved by grouping was minimal.

Wilhelms and Westby-Gibson (1961) also investigated the quest for homogeneity and discovered that any criterion or combination index for grouping reduced the total group variability by only a fifth. As would be expected, individual variability is much greater than group variability; thus, the group cannot be homogeneous because individuals are not homogeneous within themselves. Wilhelms and Westby-Gibson also offered a warning that grouping often presents

a false security about group level and, thus, less provision is made for individual differences.

A large amount of research on ability grouping is available in the literature. For the sake of economy, only major research since 1960 is presented. Of eight major research studies, five reported no significant differences (Balow, 1962; Morgenstern, 1963; Lunn, 1970; Borg, 1964; DeGrow, 1963). One was significant for homogeneous (Berkun, Swanson, and Sawyer, 1966) and one for heterogeneous (Koontz, 1961) grouping, and one reported mixed results (Goldberg, Passow, and Justman, 1966).

Of several research reviews of ability grouping, only the earliest review (Cornell, 1936) favored grouping. Later research reviews (Wilhelms and Westby-Gibson, 1961; Eckstrom, 1961; Eash, 1961; National Education Association, 1968; Esposito, 1973; Findlay and Bryan, 1975) indicated inconclusive or detrimental effects from ability grouping.

Balow (1962) addressed the issue of homogeneous grouping in a study conducted to answer two questions: (1) Are homogeneous groups really homogeneous? (2) Do homogeneous groups make greater gains in reading achievement than do heterogeneous groups? Initial investigation of reading achievement scores showed that though grouping based upon total reading scores appeared to be homogeneous, further analysis of eight subtest scores showed tremendous overlap

and very little homogeneity. The second question was tested by forming students in the sixth grade into both homogeneous and heterogeneous classes. The variance of the intelligence quotients of the two groups showed no statistically significant differences. Reading gain scores were analyzed at the end of the year, and no statistical differences were found.

Longitudinal studies by Morgenstern (1963), DeGrow (1963), and Borg (1964) resulted in no significant differences between homogeneously and heterogeneously grouped classes. Morgenstern presented a three-year study on the effects of ability grouping. Sixth-grade students were grouped either heterogeneously or homogeneously by academic ability for a three-year period. Data were collected from Stanford Achievement Test results. Data were examined for the total population of the control and experimental group for a selected group of children of low, middle, and high I.Q., and for twenty pairs of boys and girls who had been equated for age, I.Q., and school attendance. The statistical tests were not stated. The author did not find significantly different gains in academic achievement between the two groups.

In DeGrow's (1963) one-year study a random sample of students in grades four, five, and six were grouped homogeneously and heterogeneously by grade level based upon

reading scores, I.Q., and sex. In a four-year cross-sectional study, I.Q. and reading scores of sixth-grade students were analyzed. In addition, a three-year longitudinal study compared reading scores of 180 students who were enrolled in the experimental program in grades four, five, and six. Neither the one-year study, the four-year cross-section study, nor the three-year longitudinal study showed significant differences in the gains made by each group.

One of the few long-term studies on ability grouping was conducted by Borg (1964). The study included students in grades four through six over a four-year period, and two school districts were involved. District A used composite achievement test scores to group students into self-contained classrooms at the elementary level; District R assigned students into heterogeneous classes and differentiated the curriculum for individual differences through enrichment. Students in both districts were tested so that samples of students in the different ability levels could be compared. The author collected data on achievement-related dependent variables. The overriding conclusion drawn by Borg was that neither ability grouping nor heterogeneous grouping had a consistent effect upon achievement. The author suggested that the decision to use either type of grouping should be based upon considerations other than achievement.

The Lunn (1970) study investigated the streaming (homogeneous ability grouping) used in British schools. Lunn suggested that one possible explanation for the inconsistent findings in streaming was the failure to include the variable of teacher attitudes and methods. Lunn's study included measures on (1) students' scholastic attainment in reading, English, and arithmetic, (2) students' verbal and nonverbal reasoning, and (3) teachers' attitudes toward streaming. Lunn discovered that the teachers who favored ability grouping were more "knowledge centered," and emphasized discipline; the nonability-grouping teachers were more "child-centered" and more concerned with all-around pupil development. No difference was found in academic performance levels of grouped and nongrouped students.

Results favorable to homogeneous grouping were found by Berkun, Swanson, and Sawyer (1966). The effects of ability grouping on reading achievement were investigated with 1,098 students from grades three, four, and five in twenty-seven heterogeneous classes and eighteen homogeneous classes based upon reading achievement. The California Reading Test was the achievement measure. Students in the homogeneous classes achieved at a statistically significantly higher level (.4 year) than did students in heterogeneous classes. The benefits of grouping were greater for fifth-grade students and for students above average in reading level,

although below-average children also evidenced higher achievement in homogeneous classes than in heterogeneous classes.

Contrasting with the Berkun, Swanson, and Sawyer (1966) study were Koontz's (1961) results. Koontz used the Iowa Tests of Basic Skills to form homogeneous groups in arithmetic, language, and reading. Students in five fourth-grade classes formed the experimental group, and students in two fourth-grade classes were selected as the control group. Control students were matched with experimental students in each of the three achievement areas. Students were given materials and instructions on their achievement level. The analysis of data was accomplished by a treatment-by-levels analysis of variance factorial arrangement. There was a significant difference ($p < .01$) between the treatment groups in reading, language, and mathematics in favor of the control group. No significant interaction was found between treatment and level.

The most extensive study intended to evaluate ability grouping without attempting to adjust the curriculum was conducted by Goldberg, Passow, and Justman (1966). Five ability levels were identified, ranging from Group A, with I.Q.s of 130 or higher, to Group E, having I.Q.s of 99 or lower. Classes were organized to contain different combinations of the various ability levels. A total of eighty-six

classes were classified into fifteen patterns of differing combinations of I.Q. levels. Students were tested in reading, arithmetic, language arts, science, and social studies. Goldberg analyzed the data for three different effects: presence or absence of extreme group (gifted and slow) on the performance of other ability levels, narrowing the ability range and the performance of pupils, and relative position of any ability level and the performance of pupils. The findings of the study indicated that only in science did the presence of gifted pupils have a consistently upgrading effect for all other ability groups. In mathematics the presence of the slow pupils was consistently related to increased achievement of all other pupils. In the academic areas overall comparisons showed greater gains for pupils in classes that had the broadest range (four or five ability groups). With regard to position for any class with A and/or or E pupils, the A pupils always had the highest achievement and E pupils always had the lowest. For the three intermediate groups, no one position was consistently superior to any other for all ability levels in all subjects.

Cornell (1936) addressed reasons for conflicting results in research on grouping. Three reasons cited by the author were: (1) results can be evaluated only in light of conditions and purposes of experiments, which vary greatly;

(2) many of the desired results are difficult to measure; and (3) it is impossible to control all variables in an actual teaching situation. A review of several studies made by Cornell indicated that the studies had been too varied to give a true evaluation picture, but, when all aspects were considered, the results seemed to be favorable to grouping. The author did note that results seemed to depend less upon grouping than upon the philosophy of the grouping, the differences in content and method, and the techniques of the teachers.

Wilhelms and Westby-Gibson (1961) reviewed the research on grouping and offered these conclusions: (1) there was no evidence that ability grouping taken by itself increased achievement, (2) the slight edge toward ability grouping shown in many studies was minuscule, (3) the high group did not show noteworthy gains as had been purported, (4) the low group gained most from ability grouping as in the special education programs, and (5) teachers believed that grouping aided instruction.

Eckstrom (1961) limited the review of grouping studies to those involving experimental controls and reviewed those in which attempts were made to keep the content, method, and materials the same in all groups and others in which adjustments were made to allow for differences in content, method, and materials. The author stated that the former appeared

to negate the purpose of grouping. The author's summary concluded that the studies on the effectiveness of homogeneous and heterogeneous grouping, using student achievement as a measure, showed a great variety of experimental designs and no consistency in findings. Experiments that provided for different teaching methods and materials or made an effort to push bright homogeneous classes tended to favor homogeneous groups.

From his review of research Eash (1961) presented four generalizations regarding ability groupings: (1) ability grouping in itself did not produce improved achievement, (2) more recent research (after the 1920s) apparently indicated that ability grouping may have been detrimental to children in the average and lower ability groups, (3) ability grouping at an early age favored placement of children from the higher socioeconomic class in higher ability groups, and (4) ability grouping may have promoted group norms that did not foster acceptance of broad individual differences.

Findlay and Bryan (1975) drew these four conclusions in a review of the literature on ability grouping: (1) homogeneous ability grouping did not show consistent positive effects for increasing student achievement; (2) homogeneous grouping did not enhance affective development; (3) homogeneous grouping tended to group students by socioeconomic

status and, therefore, was undemocratic; and (4) when ability grouping was favorable toward achievement, substantial changes were made in teaching methods, materials, and other variables that may have had greater effects than the grouping per se.

Summary

The historical development of grading plans and grouping has been presented to provide a foundation for understanding vertical and horizontal school organization. The investigation of the research of multigrade classes on achievement has included a review of the research on two methods of vertical school organization (single-grade classes versus nongraded classes, and single-grade classes versus multigrade classes) and on one method of horizontal organization (ability grouping).

By the mid-1800s the "district" schools of the seventeenth and eighteenth centuries had evolved into the single-grade structure. Educational, social, political, and economic conditions of the time had contributed to the development of the graded structure. No sooner had gradedness become firmly established across the cities of America than did educators begin to question its effectiveness. A concern for individual differences resulted in the creation of educational plans designed to accept individual

differences and avoid nonpromotion. After a series of plans from the 1860s to the early 1900s that encouraged frequent promotions, various methods of grouping, individualization, differing rates of learning, and curriculum reorganization, the movement toward nongraded classes entered school organization in the 1930s. The movement favoring nongraded classes peaked in the 1960s, and interest diminished in the 1970s. The multigrade classroom, a common feature in the 1800s, began to reappear in the 1930s, although most multigrade plans were described in the 1960s. Though the attempts at school reorganization since the mid-1800s have been numerous, the graded structure has prevailed.

Most research on nongraded versus single-grade classes was conducted in the 1960s and early 1970s. Reviews by DiLorenzo and Slater (1965), Heathers (1969), Pavan (1973), Martin and Pavan (1976), and Heddens (1982) favored nongraded over single-grade classes. Of eight individual studies on the achievement effects of nongraded classes cited, five studies favored nongraded programs. The three longitudinal studies favored nongraded programs in all achievement areas in grades four and five (Morris, 1969), in grades one through six (Killough, 1971), and in primary (Ward, 1970). Primary students achieved at a significantly higher rate in six of fourteen subtests (Halliwell, 1963), and first-grade students scored significantly higher on two

of three achievement measures (Hillson, 1965). Achievement in reading and language favored the graded program at a significant level in intermediate grades. Two studies (Remacle, 1971; Bowman, 1971) showed mixed results. In the Remacle study nongraded sixth-grade students scored significantly higher than graded students did in five of six achievement areas; nevertheless, at the fifth grade no significant differences were found, although scores were higher for the nongraded group. Results differed at primary and intermediate grades in the Bowman study; no significant differences were observed at primary, but at the intermediate level nongraded students scored significantly higher. Thus, although the research generally favored the nongraded program, the results were not conclusive.

Most of the fourteen investigations of multigrade classes from the 1930s to the present indicated no significant differences between achievement gains of students in multigrade and single-grade classes. Four studies evidenced results that were mixed, one study favored single-grade classes, and the remaining nine studies showed no significant differences. All four of the primary-grade studies (Harvey, 1974; MacDonald and Wurster, 1974; Adair, 1978; Lincoln, 1981) showed no significant differences between single-grade and multigrade classes. Of the three studies limited to intermediate students, two (Drier, 1949; Adams,

1953) showed no significant differences and one (Foshay, 1948) reported results favoring single-grade classes.

Of the seven multigrade studies that included both primary and intermediate students, three (Knight, 1938; Chace, 1961; Way, 1969) reported statistically nonsignificant results and four showed mixed results. In the Rehwoldt and Hamilton study (1957), reading achievement favored multigrade classes at first, third, and fifth grades, and language results showed no significant differences. The Finley and Thompson (1963) study had arithmetic results that favored multigrade classes at third grade but favored single-grade boys at fifth grade, while other results were not statistically significant. Yerry's (1964) study showed no significant differences between single-grade and multigrade classes at grades two, three, and six; significant differences in grade one favored multigrade classes in arithmetic, language, and total achievement; and at grade five significant differences favored multigrade classes in reading, arithmetic, and total achievement. The Milburn (1981) results favored multigrade classes in vocabulary but found no significant differences on other tests.

Only two multigrade studies (Yerry, 1964; Milburn, 1981) considered the achievement of lower- and upper-level students, and both reported mixed results. Yerry's study favored upper-level students at primary grades but found no

differences at intermediate grades. Milburn found that students in the lower level outscored their single-grade counterparts, while students in the upper level showed no differences. None of the studies considered the criterion for placing students in the multigrade classes.

As the National Education Association (1968) suggested, given the amount and duration of interest in ability grouping, it would seem that the issue should be settled. In fact, two studies (Clarke, 1958; Wilhelms and Westby-Gibson, 1961) that analyzed the reductions of variability from ability grouping found that variability was decreased by only one-fifth. Of seven research reviews of ability grouping, only one early review (Cornell, 1936) favored ability grouping. More recent reviews (Wilhelms and Westby-Gibson, 1961; Eckstrom, 1961; Eash, 1961; National Education Association, 1968; Esposito, 1973; Findlay and Bryan, 1975) suggested that ability grouping by itself did not increase achievement. In fact, it was teaching methods, materials, and other variables that had greater effects than the grouping procedure. Of eight recent studies, five (Morgenstern, 1963; Balow, 1962; Lunn, 1970; Borg, 1964; Degrow, 1963) showed no significant differences in achievement. A study by Berkun, Swanson, and Sawyer (1966) favored homogeneous grouping, and a study by Koontz (1961) favored heterogeneous grouping. A comprehensive study by Goldberg, Passow, and

Justman (1966) reported mixed results: only in science did the presence of gifted students increase the results of other students, in mathematics the presence of slow students was consistently related to increased achievement of all other students, and in all academic areas greater gains were shown for the most heterogeneous classes. Thus, though most research showed no significant differences between heterogeneous and homogeneous groupings, the results remained inconclusive (Bolvin, 1982).

CHAPTER THREE

Methodology

The procedures discussed in this chapter are divided into five sections. The first section provides the background of multigrade classes in the school district. The second section discusses the population and describes the sampling procedures utilized. The third section describes the test instruments used in this study. The research design used in this study is explained in the fourth section, and the procedure for analysis of data is presented in the final section.

Background

The Mesa Public School District forms multigrade classes to reduce class loads at two grade levels within a single school, and in so doing reduces the number of teachers that otherwise would be required for single-grade classes at those grade levels. The multigrade classes are combined from two adjacent levels, e.g., third and fourth grades, or fourth and fifth grades. Students from the two grades are placed with one teacher but retain their

grade-level assignments and maintain their grade-specific curriculums.

During the 1981-82 school year, fifteen elementary schools operated twenty-one multigrade classes in grades two through six. The breakdown of multigrade classes follows:

<u>Multigrade Class</u>	<u>Number of Classes</u>
Second/Third	Six
Third/Fourth	Five
Fourth/Fifth	Three
Fifth/Sixth	Eight

When projection figures indicated that a multigrade class would be necessary to equalize class sizes within a school, the building principal had to make two decisions: (1) who would teach the multigrade class, and (2) what criteria would be used to place students in that class. The teachers for the multigrade classes were selected in a variety of ways, which resulted in the assignment of both (1) tenured teachers who volunteered or were asked to teach multigrade classes and (2) newly hired experienced or inexperienced teachers. The competency of individual teachers was not investigated in the study.

Investigation of student assignments revealed that achievement was a placement criterion for all schools. The

behavior of the student and the ability of the student to work independently; were other factors sometimes considered for placement. Additionally, it was discovered that the principals attempted to minimize the academic range within a given class. Based upon reading ability, the principal and/or teachers mostly placed students of average or high achievement in the multigrade classes. Three forms of placement resulted:

Placement one: A multigrade high class that included students from the upper third in academic achievement, primarily reading achievement, for both grades; e.g., high-achieving second-graders were placed with high-achieving third-graders.

Placement two: A multigrade average/high class that included students from the middle and upper thirds in academic achievement, primarily reading achievement, for both groups; e.g., average and high-achieving second-graders were placed with average and high-achieving third-graders.

Placement three: A multigrade average class that combined average students from the lower grade with average students of the upper grade; e.g., average second-graders were placed with average third-graders.

Population and Sample

Data were collected during the 1980-81 and 1981-82 school years from the Mesa Public School District in Mesa, Arizona. The district encompasses approximately 202 square miles. Students primarily are from middle-class families; 15 percent of the pupils represent minority populations. The district serves the Salt River Indian Reservation and the Pima-Papago Indian Reservation. Major categories of parent occupations include professional, small business, and service activities. The community includes a large number of retired persons.

The population of the community and the school district in 1981-82 was approximately 200,000. The total public school enrollment was 40,000, including grades kindergarten through twelve in forty-six schools. The subject sample of the study was selected from the 21,500 students enrolled in grades kindergarten through six in the thirty-two elementary schools included in the district during the 1981-82 school year.

The population of students placed in multigrade classes for the 1981-82 school year met four criteria for inclusion in the study: (1) they were assigned to multigrade classes for only the 1981-82 school year, (2) they were placed in the classes primarily based upon their achievement records, (3) they were selected at random from either the middle or

high achievement groups for their respective grades, and (4) California Achievement Test (CAT) scores from tests administered in the spring of 1981 and in the spring of 1982 would be available.

Because formation of multigrade classes in the district was based upon individual school needs, multigrade classes were not uniformly distributed across the district. One school may have had a second/third multigrade class, while another may have had a third/fourth multigrade class. The two groups of third-grade students were not in the same situation; i.e., one third-grade student group was in the higher grade of a multigrade class, and the other was in the lower grade of its combined class. The higher and lower grade assignments were identified as the level that was included for data analysis.

In addition to the multigrade class (Group 1), two other groups were formed to consider the effects of multigrade classes on student achievement. Group 2 included students enrolled in single-grade classes at the fifteen schools in the district having both single-grade and multigrade classes.

Group 3 included students enrolled in single-grade classes at the seventeen schools in the district that had no multigrade classes. This group was included as a control for selection factors other than achievement. For example,

11
if students were placed in multigrade classes based upon achievement and an unknown factor x, then the remaining students (Group 2) at those grades in single-grade classes may not have had factor x. Students in single-grade classes at other schools should have had factor x.

Over eight thousand students were selected in the original sample. Students in Groups 1 and 2 were selected from the fifteen elementary schools in the Mesa Public School District at which multigrade classes were taught. Thirteen other schools were selected from the remaining seventeen, considering achievement, mobility, ethnic distribution, and sex ratios. Frequency distributions on the California Achievement Tests with corresponding means, standard deviations, and minimum and maximum values then were found for Group 1 by placement within the grade. Students were selected randomly based upon test scores from Groups 2 and 3 to form distributions similar to those found for Group 1. The same procedures were followed for reading and mathematics scores.

Twelve separate one-way analyses of variance compared the 1981 reading scores for Groups 1, 2, and 3 within placement and grade; the same procedures were followed for mathematics scores. Analysis of variance tests applied to reading and mathematics scores indicated that differences were statistically nonsignificant. Reading mean scores and

numbers of students in each cell are presented in Table 1; mathematics data are presented in Table 2.

Sex and ethnicity for groups within placement and grade were analyzed using the contingency chi-square test for both reading and mathematics samples. For the reading sample, sex was significant, $p < .02$, for grade three, placement three by group, and $p < .04$ for grade six, placement one by group; and ethnicity was significant, $p < .02$, for grade six, placement two by group. For the mathematics sample, ethnicity was significant, $p < .02$, for grade six, placement two by group. Analysis of variance tests with 1981 reading or mathematics achievement scores were run on these significant sex and ethnicity cells; no significant differences were found for either the reading or mathematics samples.

Frequency distributions on the California Achievement Tests with corresponding means, standard deviations, and minimum and maximum value were then found for Levels 1 and 2 by placement within grades three through five. Students were selected randomly from the reading and mathematics sample to form similar distributions for the upper and lower levels. Analysis of variance tests were then run comparing the 1981 scores for the upper and lower levels; the ANOVAs run on reading and mathematics scores were found to be statistically nonsignificant. Reading mean scores and numbers of students in each cell are presented in Table 3;

Table 1
Summary of Group Within Placement and Grade on the 1981
California Achievement Tests--Total Reading

Placement and Group	Grade							
	Three		Four		Five		Six	
	N	Mean	N	Mean	N	Mean	N	Mean
Placement 1: High								
Group 1 ..	71	76.5	36	80.0	13	86.2	10	87.8
Group 2 ..	101	75.1	28	79.6	10	82.3	8	85.1
Group 3 ..	120	77.1	48	80.5	10	83.2	11	86.5
Placement 2: Average/High								
Group 1 ..	29	69.3	34	69.1	78	62.8	30	64.0
Group 2 ..	157	70.0	102	67.1	358	62.6	162	65.5
Group 3 ..	84	69.2	226	69.1	316	63.4	216	66.7
Placement 3: Average								
Group 1 ..	12	58.9	15	67.3	64	53.7	34	57.6
Group 2 ..	13	54.2	30	61.1	139	51.7	145	53.3
Group 3 ..	51	56.6	56	66.3	320	54.2	223	53.1

Table 2
Summary of Group Within Placement and Grade on the 1981
California Achievement Tests--Total Mathematics

Placement and Group	Grade							
	Three		Four		Five		Six	
	N	Mean	N	Mean	N	Mean	N	Mean
Placement 1: High								
Group 1 ..	70	71.0	36	73.8	13	86.8	10	89.7
Group 2 ..	46	72.3	18	72.7	7	78.7	4	82.8
Group 3 ..	72	73.9	40	70.6	5	92.6	4	81.3
Placement 2: Average/High								
Group 1 ..	31	66.0	33	63.3	73	59.4	30	58.7
Group 2 ..	137	66.5	95	59.1	348	59.3	154	60.9
Group 3 ..	67	65.2	204	63.0	286	60.0	209	60.7
Placement 3: Average								
Group 1 ..	8	59.4	11	53.3	60	52.0	33	50.8
Group 2 ..	9	54.1	22	52.5	125	49.1	139	51.6
Group 3 ..	25	58.0	46	56.1	256	50.1	204	54.6

Table 3

Summary of Level Within Placement and Group on the 1981
California Achievement Tests--Total Reading

Placement and Level	Grade					
	Three		Four		Five	
	N	Mean	N	Mean	N	Mean
Placement 1: High						
Level 1	30	76.0	not enough cases		not enough cases	
Level 2	41	76.9				
Placement 2: Average/High						
Level 1	12	74.0	not enough cases		37	65.6
Level 2	17	65.9			41	60.3
Placement 3: Average						
Level 1	not enough cases		7	68.6	14	43.6
Level 2			8	66.1	16	43.5

mathematics data are presented in Table 4. For some cells there were not enough cases to run the analyses.

Analysis of variance tests were run on the reading and mathematics samples using a 3 (group) by 4 (grade) by 3 (placement) factorial arrangement to ensure that the original findings of no statistically significant difference among groups within placement and grade held true when combined over grades and place. The analysis of reading scores reported in Table 5 shows that group was nonsignificant at the .05 level, but both placement and grade were significant at the .001 level. Also, group by grade and group by placement interactions were nonsignificant, but grade by placement was significant ($p < .001$).

The analysis of variance data on mathematics scores are shown in Table 6. Group was statistically nonsignificant, but both grade and placement were significant at the .001 level. The two-way interactions of group by grade and group by placement were nonsignificant, but grade by placement was significant ($p < .001$).

Both mathematics and reading samples indicated an interaction of placement and grade. In addition, the sampling procedures resulted in unequal sample sizes, which in a factorial arrangement yield biased results (Ferguson, 1976). Because of these confounding factors, further analyses used the one-way analysis of variance run for group

Table 4

Summary of Level Within Placement and Group on the 1981
California Achievement Tests--Total Mathematics

Placement and Level	Grade					
	Three		Four		Five	
	N	Mean	N	Mean	N	Mean
Placement 1: High						
Level 1	30	68.0	not enough cases		not enough cases	
Level 2	25	69.0				
Placement 2: Average/High						
Level 1	8	73.1	not enough cases		33	60.9
Level 2	6	66.3			40	58.2
Placement 3: Average						
Level 1	not enough cases		6	53.2	15	44.8
Level 2			5	53.4	15	42.1

Table 5
Summary of Analysis of Variance of Reading Achievement
by Group, Grade, and Placement

Source	SS	DF	MS	F	p <
Group	965.05	2	482.52	2.69	.066
Grade	15395.40	3	5131.80	28.61	.001*
Place	123953.82	2	61976.91	345.48	.001*
Group x grade	380.11	6	63.35	.35	.999
Group x place	583.83	4	145.96	.81	.999
Grade x place	12817.00	6	2136.17	11.91	.001*
Group x grade x place	1031.20	12	85.93	.48	.999
Residual	596307.87	3324	179.40		
Total	825634.88	3359	245.80		

*Significant at $p < .001$.

Table 6
Summary of Analysis of Variance of Mathematics Achievement
by Group, Grade, and Placement

Source	SS	DF	MS	F	p <
Group	643.58	2	321.79	1.74	.173
Grade	6186.91	3	2062.31	11.17	.001*
Place	81378.85	2	40689.42	220.33	.001*
Group x grade	968.96	6	161.49	.87	.999
Group x place	313.69	4	78.42	.43	.999
Grade x place	11119.08	6	1853.18	10.04	.001*
Group x grade x place	2180.64	12	181.72	.984	.999
Residual	534826.52	2896	184.68		
Total	684801.13	2931	233.64		

*Significant at $p < .001$.

within placement and grade. Thus, the interaction effects, group by grade and group by placement, could not be isolated statistically with this type of analysis, but inferences could be made about differential effects.

Instrumentation

The instrument chosen for assessing reading and mathematics achievement was the California Achievement Tests, Form C. At the time of this study the test was the state achievement test administered to all students in grades one through twelve in April of each school year. The spring 1981 test results and the spring 1982 test results were used for the study. Different levels of the test were administered to different grade levels, but the normal curve equivalencies allowed for using different levels of a test for analysis.

Investigations of the California Achievement Tests (CAT) have revealed them to be both reliable and valid. On a scale of 0 to 1, the CAT yielded coefficients ranging from .72 to .98 on internal consistency and .35 to .95 on consistency over time, but the .35 coefficient was unrepresentative of the other measures of consistency over time. Most of the coefficients were greater than .60, which was considered an acceptable score on reliability. The Short Form Test of Academic Aptitude (SFTAA) was

used to validate the CAT. Scores on these tests showed correlations ranging from .74 to .86.

Design

The problem in this study was to determine the effects of multigrade classes on student achievement in reading and mathematics. The research questions addressed were:

1. Is there a significant difference among the sample means of three grouping patterns as measured by reading and mathematics achievement tests?
2. Is there a significant interaction among the sample means of three grouping patterns and four different grades as measured by reading and mathematics achievement tests?
3. Is there a significant interaction among the sample means of three groups and three placement types as measured by reading and mathematics achievement tests?
4. Is there a significant difference among the sample means of two levels as measured by reading and mathematics achievement tests?
5. Is there a significant interaction among the sample means of three placement types, two levels, and three grades as measured by reading and mathematics achievement tests?

The statistical design selected to answer these questions was the one-way analysis of variance using only a posttest. If significant differences were found, Fisher's

LSD was used to compare the means. The one-way analysis of variance was selected because it was run on the smallest cell that was nonsignificant and because the one-way analysis of variance yielded an unbiased test with unequal sample size. Administration of a pretest was not essential to true experimental design. Since this was a quasi-experimental design, tests were run on 1981 CAT data to ensure that the samples used in the group comparisons could be considered as coming from the same population (Isaac and Michael, 1972).

Analysis of Data

Data used in the statistical analysis for this study came from the California Achievement Tests. Data processing was completed on the Honeywell 6000 series computer at Mesa Public Schools. The Statistical Program for the Social Sciences (SPSS) was used to analyze data.

The SPSS computed means, standard deviations, and the standard errors of the means. Analyses included a one-way analysis of variance on achievement scores and contingency chi-square tests on demographic data. Each F-ratio was then tested for significance at the .05 level. Fisher's LSD was run on achievement scores that were significant.

CHAPTER FOUR

Findings

The purpose of this study was to investigate the effects of multigrade classes on student achievement in reading and mathematics. The first research question centered upon the differences among the means of the three groups.

Question 1: Is there a significant difference among the sample means of three grouping patterns as measured by reading and mathematics achievement tests?

An analysis of reading data from the California Achievement Tests for the three groups within placement and grade is provided in Tables 7 through 18. The F-test from the analysis of variance computed to answer this question was statistically significant only at grade four for placement one, which included high-achieving students. An inspection of the F-ratio found in Table 10 showed significance ($p < .01$). Results were statistically nonsignificant for all other grades and placements.

Table 7

Summary of Group Within Grade 3 and Placement 1 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	776.92	2	388.46	2.42	.09
Within groups	46310.95	289	160.25		
Total	47087.88	291			

Table 8

Summary of Group Within Grade 3 and Placement 2 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	214.16	2	107.08	.57	.57
Within groups	50288.63	267	188.35		
Total	50502.78	269			

Table 9

Summary of Group Within Grade 3 and Placement 3 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	307.83	2	153.92	.99	.38
Within groups	11389.17	73	156.02		
Total	11697.00	75			

Table 10

Summary of Group Within Grade 4 and Placement 1 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	1490.20	2	745.10	4.40	.01*
Within groups	18453.66	109	169.30		
Total	19943.86	111			

*Significant at $p < .05$

Table 11

Summary of Group Within Grade 4 and Placement 2 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	861.56	2	430.78	1.95	.14
Within groups	79348.77	359	221.03		
Total	80210.33	361			

Table 12

Summary of Group Within Grade 4 and Placement 3 on the
1982 California Achievement Tests--Total reading

Source	SS	DF	MS	F	p <
Between groups	202.39	2	101.20	.51	.61
Within groups	19487.85	98	198.86		
Total	19690.24	100			

Table 13

Summary of Group Within Grade 5 and Placement 1 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	224.07	2	112.04	.727	.50
Within groups	4624.17	30	154.14		
Total	4848.24	32			

Table 14

Summary of Group Within Grade 5 and Placement 2 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	1081.53	2	540.77	2.38	.09
Within groups	170413.53	749	227.52		
Total	171495.06	751			

Table 15

Summary of Group Within Grade 5 and Placement 3 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	1654.34	2	827.17	2.58	.08
Within groups	167009.77	520	321.17		
Total	168664.11	522			

Table 16

Summary of Group Within Grade 6 and Placement 1 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	23.58	2	11.79	.11	.89
Within groups	2864.28	26	110.16		
Total	2887.86	28			

Table 17

Summary of Group Within Grade 6 and Placement 2 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	563.72	2	281.86	1.49	.23
Within groups	76542.11	405	188.99		
Total	77105.83	407			

Table 18

Summary of Group Within Grade 6 and Placement 3 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	155.14	2	77.57	.374	.69
Within groups	82813.47	399	207.55		
Total	82968.61	401			

The results of the least significant difference procedure are presented in Table 19. The procedure showed that for grade four and placement one, Group 2 and Group 3 were not significantly different from each other, but both were significantly different from Group 1.

Table 19
Least Significant Difference Procedure for Groups Within
Grade 4 and Placement 1 on the 1982 California
Achievement Tests--Total Reading

Group 1 Mean	Group 3 Mean	Group 2 Mean
78.7	<u>71.6</u>	<u>69.9</u>

Note: The dashed underline indicates homogeneous subsets.

Mathematics data from the California Achievement Tests for the three groups within placement and grade appear in Tables 20 through 31. The F-test from the analysis of variance calculated to answer this question was significant for five of the twelve analyses. The F-ratio of 3.90 shown in Table 20 indicated significance for grade three and placement one, high-achieving students, at the $p < .02$ level.

Table 20

Summary of Group Within Grade 3 and Placement 1 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	1271.77	2	635.89	3.90	.02*
Within groups	50379.18	185	164.21		
Total	31650.95	187			

*Significant at $p < .05$.

Table 21

Summary of Group Within Grade 3 and Placement 2 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	99.37	2	49.68	.22	.81
Within groups	53251.08	232	229.53		
Total	53350.45	234			

Table 22

Summary of Group Within Grade 3 and Placement 3 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	1581.78	2	790.89	4.87	.01*
Within groups	6332.22	39	162.36		
Total	7914.00	41			

*Significant at $p < .05$.

Table 23

Summary of Group Within Grade 4 and Placement 1 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	389.30	2	194.65	.99	.38
Within groups	17980.52	91	197.59		
Total	18369.83	93			

Table 24

Summary of Group Within Grade 4 and Placement 2 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	1058.31	2	529.16	1.49	.22
Within groups	116539.36	329	354.22		
Total	117597.67	331			

Table 25

Summary of Group Within Grade 4 and Placement 3 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	158.55	2	79.27	.359	.70
Within groups	16773.20	76	220.70		
Total	16931.75	78			

Table 26

Summary of Group Within Grade 5 and Placement 1 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	530.01	2	265.01	2.28	.12
Within groups	2556.14	22	116.19		
Total	3086.16	24			

Table 27

Summary of Group Within Grade 5 and Placement 2 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	1333.63	2	666.81	2.65	.07
Within groups	177487.88	706	251.40		
Total	178821.50	708			

Table 28

Summary of Group Within Grade 5 and Placement 3 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	2004.75	2	1002.38	3.72	.03*
Within groups	118174.64	438	269.81		
Total	120179.39	440			

*Significant at $p < .05$.

Table 29

Summary of Group Within Grade 6 and Placement 1 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	129.21	2	64.61	.49	.63
Within groups	1961.90	15	130.79		
Total	2091.11	17			

Table 30

Summary of Group Within Grade 6 and Placement 2 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	1404.39	2	702.20	3.09	.05*
Within groups	88739.86	390	227.54		
Total	90144.25	392			

*Significant at $p < .05$.

Table 31

Summary of Group Within Grade 6 and Placement 3 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	4936.02	2	2468.01	11.66	.001*
Within groups	78983.22	373	211.75		
Total	83919.23	375			

*Significant at $p < .05$.

The F-ratio for 4.87 shown in Table 22 indicated significance for grade three and placement three, average-achieving students, at the $p < .01$ level. The result of the analysis of variance test for grade five and placement three, average-achieving students, yielded an F-ratio of 3.72, which was significant at the $p < .03$ level (see Table 28). Results of the analysis of variance for grade six and placement two, average/high-achieving students, provided an F-ratio of 3.09, which was significant at the $p < .05$ level (see Table 30). The F-ratio of 11.66 shown in Table 31 reported significance for grade six and placement three, average-achieving students, at the $p < .001$ level.

The results of the least significant difference procedures for mathematics are summarized in Tables 32 through 36. For grade three and placement one, shown in Table 32, Group 3 and Group 2 and also Group 2 and Group 1 were not significantly different from each other; Groups 1 and 3 were significantly different from each other. In Table 33, which presents data for grade three and placement three, Group 3 and Group 1, as well as Group 2 and Group 1, were not significantly different from each other; Groups 2 and 3 were significantly different.

The results of the least significant difference procedure for grade five and placement three are presented in Table 34. Group 2 was not statistically significantly

Table 32

Least Significant Difference Procedure for Groups Within
Grade 3 and Placement 1 on the 1982 California
Achievement Tests--Total Mathematics

Group 3 Mean	Group 2 Mean	Group 1 Mean
<u>76.2</u>	<u>71.6</u>	70.5

Note: The dashed underlines indicate homogeneous subsets.

Table 33

Least Significant Difference Procedure for Groups Within
Grade 3 and Placement 3 on the 1982 California
Achievement Tests--Total Mathematics

Group 3 Mean	Group 1 Mean	Group 2 Mean
<u>64.6</u>	<u>57.5</u>	49.4

Note: The dashed underlines indicate homogeneous subsets.

Table 34

Least Significant Difference Procedure for Groups Within
Grade 5 and Placement 3 on the 1982 California
Achievement Tests--Total Mathematics

Group 3 Mean	Group 2 Mean	Group 1 Mean
55.3	<u>51.3</u>	<u>50.3</u>

Note: The dashed underline indicates homogeneous subsets.

Table 35

Least Significant Difference Procedure for Groups Within
Grade 6 and Placement 2 on the 1982 California
Achievement Tests--Total Mathematics

Group 1 Mean	Group 3 Mean	Group 2 Mean
73.3	67.8	65.9

Note: The dashed underlines indicate homogeneous subsets.

Table 36

Least Significant Difference Procedure for Groups Within
Grade 6 and Placement 3 on the 1982 California
Achievement Tests--Total Mathematics

Group 3 Mean	Group 2 Mean	Group 1 Mean
60.4	54.0	50.8

Note: The dashed underline indicates homogeneous subsets.

different from Group 1, but both were significantly different from Group 3. For grade six and placement two, given in Table 35, Group 1 and Group 3 and also Groups 3 and 2 were not significantly different from each other; Groups 1 and 2 were significantly different. As shown in Table 36, which includes data on grade six and placement three, Group 2 and Group 1 were not significantly different from each other; both were significantly different from Group 3.

Question 2: Is there a significant interaction among the sample means of three grouping patterns and four different grades as measured by reading and mathematics achievement tests?

The data in Table 37 showed that for reading the groups were not statistically significant for grades three, five, and six. At grade four, groups were significantly different for one of three placements.

The mathematics data presented in Table 38 indicated statistically significant results within each grade except grade four; no two grades had the same results from the one-way analysis of variance tests.

Question 3: Is there a significant interaction among the sample means of three groups and three placement types as measured by reading and mathematics achievement tests?

Table 37

Summary of Group Within Placement and Grade on the 1982
California Achievement Tests--Total Reading

Placement and Group	Grade							
	Three		Four		Five		Six	
	N	Mean	N	Mean	N	Mean	N	Mean
Placement 1: High								
Group 1 ..	71	76.2	36	78.7*	13	83.7	10	83.7
Group 2 ..	101	72.1	28	69.9*	10	79.5	8	81.5
Group 3 ..	120	74.8	48	71.6*	10	86.1	11	83.3
Placement 2: Average/High								
Group 1 ..	29	67.8	34	67.3	78	67.5	30	66.4
Group 2 ..	157	65.7	102	62.3	358	64.0	162	63.5
Group 3 ..	84	67.4	226	65.1	316	65.8	216	65.9
Placement 3: Average								
Group 1 ..	12	51.3	15	57.3	64	52.9	34	54.3
Group 2 ..	13	54.5	30	61.3	139	53.2	145	52.1
Group 3 ..	51	56.7	56	61.3	320	56.7	223	52.9

*Significant at $p < .05$.

Table 38

Summary of Group Within Placement and Grade on the 1982
California Achievement Tests--Total Mathematics

Placement and Group ^a	Grade							
	Three		Four		Five		Six	
	N	Mean	N	Mean	N	Mean	N	Mean
Placement 1: High								
Group 1 ..	70	70.5*	36	79.3	13	85.5	10	84.4
Group 2 ..	46	71.6*	18	73.7	7	76.4	4	90.3
Group 3 ..	72	76.2*	40	76.6	5	88.6	4	89.3
Placement 2: Average/High								
Group 1 ..	31	63.4	33	68.6	73	64.8	30	73.3*
Group 2 ..	137	64.9	95	62.7	348	61.5	154	65.9*
Group 3 ..	67	63.6	204	65.9	288	64.1	209	67.8*
Placement 3: Average								
Group 1 ..	8	57.5*	11	55.0	60	50.3*	33	50.8*
Group 2 ..	9	49.4*	22	58.9	125	51.3*	139	54.0*
Group 3 ..	25	64.6*	46	59.2	256	55.3*	204	60.4*

*Significant at $p < .05$.

The results of the reading data, listed in Table 37, indicated that groups were not statistically significant for the average/high placement and the average placement. Groups were significantly different for the high placement at grade four.

The mathematics data, presented in Table 38, showed that groups were significantly different for grade three within the high placement, for grade six within the average/high placement, and for grades three, five, and six within the average placement.

Question 4: Is there a significant difference among the sample means of two levels as measured by reading and mathematics achievement tests?

An analysis of reading data from the California Achievement Tests for level within placement and grade is detailed in Tables 39 through 43 for cells that had sufficient cases to run the analyses. Levels for all grades and placements were statistically nonsignificant for reading.

Mathematics data from the California Achievement Tests for level within placement and grade are summarized in Tables 44 through 48. The F-test from the analysis of variance was statistically significant for grade five and placement three, which included average-achieving students. Levels at other grades and placements were not significant for mathematics.

Table 39

Summary of Level Within Grade 3 and Placement 1 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	2.32	1	2.32	.02	.87
Within groups	9697.51	69	140.54		
Total	9699.83	70			

Table 40

Summary of Level Within Grade 3 and Placement 2 on the
1982 California Achievement Tests--Total Reading

Source	SS	LF	MS	F	p <
Between groups	128.53	1	128.53	.58	.46
Within groups	5979.61	27	221.47		
Total	6108.14	28			

Table 41

Summary of Level Within Grade 4 and Placement 3 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	457.62	1	457.62	3.23	.09
Within groups	1843.71	13	141.82		
Total	2301.33	14			

Table 42

Summary of Level Within Grade 5 and Placement 2 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	1.87	1	1.87	.01	.89
Within groups	16973.62	76	223.34		
Total	16975.48	77			

Table 43

Summary of Level Within Grade 5 and Placement 3 on the
1982 California Achievement Tests--Total Reading

Source	SS	DF	MS	F	p <
Between groups	523.72	1	523.72	2.45	.13
Within groups	5985.75	28	213.78		
Total	6509.47	29			

Table 44

Summary of Level Within Grade 3 and Placement 1 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	419.53	1	419.53	3.36	.07
Within groups	6609.81	53	124.71		
Total	7028.84				

Table 45

Summary of Level Within Grade 3 and Placement 2 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	53.72	1	53.72	.37	.56
Within groups	1749.71	12	145.81		
Total	1803.43	13			

Table 46

Summary of Level Within Grade 4 and Placement 3 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	1.47	1	1.47	.01	.90
Within groups	2060.53	9	228.95		
Total	2062.00	10			

Table 47

Summary of Level Within Grade 5 and Placement 2 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	148.83	1	148.83	.50	.49
Within groups	21197.86	71	298.56		
Total	21346.69	72			

Table 48

Summary of Level Within Grade 5 and Placement 3 on the
1982 California Achievement Tests--Total Mathematics

Source	SS	DF	MS	F	p <
Between groups	57.63	1	57.63	5.04	.03*
Within groups	2984.67	28	106.60		
Total	3522.30	29			

*Significant at $p < .05$.

Question 5: Is there a significant interaction among the sample means of three placement types, two levels, and three grades as measured by reading and mathematics achievement tests?

The results of the reading data analyses, presented in Table 49, indicated that levels were statistically non-significant for all of the grades and placements that had sufficient cases for analysis. The mathematics data, set forth in Table 50, showed that levels were nonsignificant for grades three and four and for placements one and two. For grade five and placement three, average-achieving students, the means were significantly different.

Table 49
Summary of Level on the 1982 California Achievement
Tests--Total Reading

Placement and Level	Grade					
	Three		Four		Five	
	N	Mean	N	Mean	N	Mean
Placement 1: High						
Level 1	30	76.0	not enough cases		not enough cases	
Level 2	41	76.4				
Placement 2: Average/High						
Level 1	12	70.3	not enough cases		37	67.3
Level 2	17	66.1			41	67.6
Placement 3: Average						
Level 1	not enough cases		7	51.4	14	50.0
Level 2			8	62.5	16	41.6

Table 50
Summary of Level on the 1982 California Achievement
Tests--Total Mathematics

Placement and Level	Grade					
	Three		Four		Five	
	N	Mean	N	Mean	N	Mean
Placement 1: High						
Level 1	30	72.5	not enough cases		not enough cases	
Level 2	25	66.9				
Placement 2: Average/High						
Level 1	8	67.1	not enough cases		33	66.4
Level 2	6	63.2			40	63.5
Placement 3: Average						
Level 1	not enough cases		6	55.3	15	47.5*
Level 2			5	54.6	15	39.1*

*Significant at $p < .05$.

CHAPTER FIVE
Summary, Discussion, Conclusions, and
Recommendations

This chapter presents a summary of the study. The summary is followed by the discussion, conclusions, and recommendations.

Summary

The purpose of this study was to investigate the effects of multigrade classes on student achievement in reading and mathematics. In addition, the study considered the effects of (1) grade, (2) placement--high, average/high, or average ability grouping, and (3) level--the upper or lower grade in a multigrade class.

The sample included 3,360 students in third, fourth, fifth, and sixth grades at the Mesa Public Schools in Mesa, Arizona. Students in multigrade classes, the remaining students in single-grade classes at the schools having multigrade classes, and students in single-grade classes at schools not having multigrade classes constituted the three groups studied. All students were tested with the California Achievement Tests in spring 1981 and spring 1982.

A posttest-only design was utilized. The results of the California Achievement Test, Total Reading and Total Mathematics, were used to determine the answers to the five research questions asked in this study. The one-way analysis of variance test and the least significant difference test were used to analyze the data at the $p < .05$ level of significance.

The research questions and the findings to the five questions were:

Question 1: Is there a significant difference among the sample means of three grouping patterns as measured by reading and mathematics achievement tests?

In reading achievement, the multigrade classes were significantly different from the single-grade classes at grade four, placement one, which included high-achieving students. The multigrade classes had the highest mean, which was significantly different from that of the other groups. In mathematics achievement, multigrade classes were significantly different from the single-grade classes for five of twelve analyses. Differences existed for grade three, placement one, high-achieving students, and for placement three, average-achieving students; for grade five, placement three, average-achieving students; for grade six, placement two, high/average-achieving students, and for

placement three, average-achieving students. Group 1, multigrade classes, had the highest means at grade six and placement two, average/high-achievement students; Group 3, single-grade classes at schools not having multigrade classes, had the highest means for the remaining four analyses. Based upon the test for least significant differences, Group 3 was significantly higher than the other groups only at grade five, placement three, and at grade six, placement three.

Question 2: Is there a significant interaction among the sample means of three grouping patterns and four different grades as measured by reading and mathematics achievement tests?

In reading achievement, the groups were not significantly different for grades three, five, and six. At grade four, statistical differences occurred that favored the multigrade classes.

In mathematics achievement, the converse was true; grades three, five, and six had significant differences within the grade, but grade four did not. At grade four, differences favored multigrade classes. For grades three and five, differences favored single-grade classes in schools not having multigrade classes. For grade six, the average/high placement favored multigrade classes, but the

average placement favored single-grade classes in schools not having multigrade classes.

Question 3: Is there a significant interaction among the sample means of three groups and three placement types as measured by reading and mathematics achievement tests?

In reading achievement, groups were not significantly different for placement two, average/high-achieving students, and for placement three, average-achieving students. Statistical significance for high-achieving students favored the multigrade classes. In mathematics, achievement groups were significantly different for all three placement types. Statistical significance favored the single-grade classes for high- and average-achieving students; multigrade classes were favored for average/high-achieving students.

Question 4: Is there a significant difference among the sample means of two levels as measured by reading and mathematics achievement tests?

In reading, for cells that had sufficient cases to analyze achievement, level was nonsignificant. In mathematics, for cells that had sufficient cases for analysis, level was nonsignificant except for grade five and placement three, average-achieving students.

Question 5: Is there a significant interaction among the sample means of three placement types, two

levels, and three grades as measured by reading and mathematics achievement tests?

In reading achievement, level was nonsignificant for all placement types and grades. In mathematics achievement, level was nonsignificant for all grades and levels, except grade five and placement three, average-achieving students. The difference favored the average-achieving fifth-graders in a fifth/sixth multigrade class.

Discussion

This study was based upon five research questions that examined the effects of multigrade classes on student reading and mathematics achievement. These questions and conclusions relevant to answering them are presented below.

Question 1: Is there a significant difference among the sample means of three grouping patterns as measured by reading and mathematics achievement tests?

For the twelve analyses comparing groups on reading achievement, only one analysis showed a significant difference: high-achieving students in grade four in multigrade classes scored significantly higher than high-achieving fourth-grade students in single-grade classes. At other grades and placements, no significant differences among the groups were found. Of the twelve analyses, eight favored (were higher, but not significantly higher than) multigrade

classes. Therefore, in reading achievement, students in multigrade classes tended to achieve as well as or better than students in single-grade classes.

For the twelve analyses comparing groups on mathematics achievement, five analyses showed significant differences: high-achieving third-graders in single-grade classes scored significantly higher than high-achieving third-graders in multigrade classes; average-achieving third-, fifth-, and sixth-graders in single-grade classes not in multigrade schools outscored average third-, fifth-, and sixth-graders in multigrade classes; average/high-achieving sixth-graders in multigrade classes scored higher than average/high-achieving sixth-graders in single-grade classes. Of the twelve analyses, four favored multigrade classes and eight favored single-grade classes. Therefore, students in multigrade classes did not tend to score as well as students in single-grade classes. Also, students in multigrade classes did not score as well in mathematics as in reading achievement.

Question 2: Is there a significant interaction among the sample means of three grouping patterns and four different grades as measured by reading and mathematics achievement tests?

Inspection of the twelve one-way analyses of variance in reading indicated that student achievement was

significantly different within grade four and not significantly different within grades three, five, and six. Only at grade six did multigrade classes achieve higher scores at all three placement types.

Inspection of the twelve one-way analyses of variance in mathematics indicated that student achievement was different for all grades. Significant differences were noted for grades three, five, and six.

Question 3: Is there a significant interaction among the sample means of three groups and three placement types as measured by reading and mathematics achievement tests?

In reading achievement, groups were not significantly different for average/high-achieving students or for average-achieving students. At the high placement, multigrade classes scored highest for three of four grades; at the average/high placement, multigrade classes scored highest for all four grades; but at the average placement, multigrade students scored highest for only one grade.

In mathematics achievement, significant differences were noted for the three placement types. At the high placement, multigrade classes scored highest for one grade only; at the average/high placement, multigrade classes scored highest for three of four grades; but at the average placement, multigrade classes did not score highest for any

grade. Therefore, in reading, the multigrade classes scored highest at the high and average/high placement, and in mathematics they scored highest at the average/high placement.

Question 4: Is there a significant difference among the sample means of two levels as measured by reading and mathematics achievement tests?

In reading achievement, level was nonsignificant. In mathematics achievement, levels were significantly different for grade five, placement three, with the means favoring the lower level. Levels did not tend to be a significant factor for reading or mathematics.

Question 5: Is there a significant interaction among the sample means of three placement types, two levels, and three grades as measured by reading and mathematics achievement tests?

In reading achievement, level was nonsignificant for all placement types and grades for the five cells that had sufficient cases for analysis. In reading, neither level was favored consistently. In mathematics achievement, levels were significantly different for grade five, placement three. Thus, the average fifth-grade students in a fifth/sixth multigrade class scored significantly higher than the average fifth-grade students in a fourth/fifth multigrade class. The lower level was favored in all of the five cells analyzed for mathematics.

Conclusions

The conclusions supported by this investigation follow.

1. No significant differences were found among the three groups in reading except for grade four, placement one, which was significantly higher for multigrade classes. Significant differences among the three groups in mathematics existed for grade three, placements one and three; for grade five, placement three; and for grade six, placement three, which were all significantly higher for single-grade classes. For grade six, placement two, the multigrade classes were significantly higher.

2. Significant differences were found among groups within grade four for reading, which were higher for multigrade classes, and within grades three, five, which were higher for single-grade classes. Grade six results were mixed, with one placement significantly higher for multigrade and the other for single-grade classes.

3. A significant difference was found among groups within placement one, high-achieving students, for reading, which was higher for multigrade classes. In mathematics, significant differences were found for high-achieving students and average-achieving students that were higher for single-grade classes. The significant differences were higher for multigrade students at the average/high placement.

4. No significant differences were found between levels for reading achievement. A significant difference existed between levels for mathematics achievement in grade five and placement three, average-achieving students.

5. No significant differences were found between levels for any grade or any placement for reading achievement. A significant difference existed between levels for mathematics achievement, which favored the average-achieving fifth-grade students in a fifth/sixth multigrade class.

Implications for Practitioners

1. Multigrade classes did not appear to affect reading achievement negatively.

2. Student mathematics achievement might be negatively affected by placement in multigrade classes. If using multigrade classes, teachers should pay particular attention to mathematics instruction.

3. The effects of multigrade classes at specific grades were inconclusive, though mathematics appeared to have been affected more negatively than was reading.

4. If using multigrade classes, the average/high placement appeared to be best for all grades for reading and for grades four, five, and six for mathematics.

5. For multigrade classes, level did not appear to affect reading achievement; level did have some

effect on mathematics achievement, which favored the lower level.

Recommendations for Researchers

1. More research is needed on multigrade classes to determine the achievement effects on grade.

2. Multigrade research studies need to consider the criteria for student placement in multigrade classes. Precise definition of criteria is advised.

3. Research studies need to be completed on the effect of multigrade classes on level--the upper or lower grade in the multigrade class.

4. Complete follow-up studies need to be completed on multigrade classes to determine whether or not reported changes in achievement continue to exist during the subsequent year.

5. Controlled research studies need to be completed on the effect of multigrade classes on student, parent, and teacher attitudes.

6. Research studies need to be conducted to determine whether or not curriculum changes need to be made for multigrade classes.

7. Instructional techniques in multigrade classes need to be identified and analyzed.

8. Alternatives to the research design and statistical techniques used in this study should be considered in any future research.

9. Research should be conducted to consider the criteria for teacher assignment to teach the multigrade class.

BIBLIOGRAPHY

- Adair, James Hartley. 1978. An attitude and achievement comparison between kindergarten and first grade children in multi and single grade classes (Doctoral dissertation, Boston College). Dissertation Abstracts International 39:659A-660A.
- Adams, Joseph J. 1953. Achievement and social adjustment of pupils in combination class enrolling pupils of more than one grade level. Journal of Educational Research 47 (October):151-55.
- Anderson, Richard C. 1962. The case for nongraded homogeneous grouping. Elementary School Journal 62 (January):193-97.
- Balow, Irving H. 1962. Does homogeneous grouping give homogeneous groups? Elementary School Journal 63 (October):28-32.
- Barr, R., and Dreeben, R. 1977. Instruction in classrooms. In Review of research in education 5, 1977. Itasca, Ill.: F. E. Peacock.
- Beggs, David W., III, and Buffie, Edward G. 1967. Non-graded schools in action. Bloomington: Indiana University Press.
- Berkun, Mitchell M.; Swanson, Lloyd W.; and Sawyer, David M. 1966. An experiment on homogeneous grouping for reading in elementary classes. Journal of Educational Research 59 (May-June):413-14.
- Bidwell, C. E., and Kasarda, J. D. 1975. School district organization and student achievement. American Sociological Review 40 (January):55-70.
- Bienvenu, Harold J., and Martyn, Kenneth A. 1955. Why few combination classes? American School Board Journal 130 (April):33-34.

- Bolvin, John O. 1982. Classroom organization. In Encyclopedia of educational research. 5th ed. New York: The Free Press.
- Borg, Walter R. 1964. An evaluation of ability grouping. Co-operative Research Project No. 577, Utah State University.
- Bowman, Betty Lynch. 1971. A comparison of pupil achievement and attitude in a graded school, 1968-69, 1969-70 school years (Doctoral dissertation, University of North Carolina at Chapel Hill). University Microfilms International 32:660A.
- Brown, B. Frank. 1965. The appropriate placement school: A sophisticated nongraded curriculum. West Nyack, N.Y.: Parker Publishing Company, Inc.
- But they can't close MY child's school! 1980. Changing Times, November, pp. 56-58.
- Butts, R. Freeman, and Cremins, Lawrence. 1953. A history of education in American culture. New York: Henry Holt and Company.
- Carbone, Robert F. 1961. Comparison of graded and non-graded elementary schools. Elementary School Journal 62 (November):82-88.
- California achievement tests technical bulletin 1. 1979. Monterey-Calif.: CTB/McGraw-Hill.
- Chace, Earl Stanley. 1961. An analysis of some effects of multiple-grade grouping in an elementary school. Unpublished doctoral thesis, University of Tennessee.
- Chanan, Jebriel, ed. 1970. Streaming and the primary teacher. London: National Foundation for Educational Research in England and Wales.
- Clarke, S. C. T. 1958. The effects of grouping on variability in achievement at the grade III level. In Grouping in education. New York: John Wiley & Sons, Inc.
- Cornell, Ethel. 1936. Effects of ability grouping determinable from published studies. In The grouping of pupils: Yearbook of the National Society for the Study of Education, Part I. Edited by G. M. Whipple. Bloomington, Ill.: Public School Publishing Company.

- Cubberly, Ellwood P. 1936. Public education in the United States. Boston: Houghton Mifflin Company.
- Dawson, M., ed. 1961. Elementary school organization. The Elementary School Principal 41 (December):77-93.
- DeGrow, Gerald Samuel. 1963. A study of the effects of the use of vertical reading ability groupings for reading classes as compared with heterogeneous groupings in grades four, five, and six in the Port Huron area public schools of Michigan over a three year period (Doctoral dissertation, University of Michigan). Dissertation Abstracts International 24:3166-67.
- DiLorenzo, Louis T., and Salter, Ruth. 1965. Co-operative research on the nongraded primary. Elementary School Journal 65 (February):269-77.
- Drier, William H. 1949. The differential achievement of rural graded and ungraded school pupils. Journal of Educational Research 43:175-85.
- Dufay, Frank R. 1966. Ungrading the elementary school. West Nyack, N.Y.: Parker Publishing Company, Inc.
- Eash, Maurice. 1961. Grouping: What have we learned? Educational Leadership 18 (April):429-34.
- Eckstrom, Ruth B. 1961. Experimental studies of homogeneous grouping: A critical review. The School Review 59 (Summer):216-26.
- Esposito, Dominick, 1973. Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. Review of Educational Research 43 (Spring):163-79.
- Ferguson, George. 1976. Statistical analysis in psychology and education. New York: McGraw-Hill Book Company.
- Findlay, W. G., and Bryan, M. M. 1975. The pros and cons of ability grouping. Bloomington, Ind.: Phi Delta Kapp Educational Foundation.
- Finley, Carmen J., and Thompson, Jack M. 1963. Comparison of the achievement of multi-graded and single-graded rural elementary school children. Journal of Educational Research 56 (May-June):471-75.

- Foshay, Arthur. 1948. Interage grouping in the elementary school. Unpublished doctoral thesis, Columbia University, Teachers College.
- Franklin, Marian Pope. 1967a. Multigrading in elementary education. Childhood Education 43 (May):513-15.
- , 1967b. School organization: Theory and practice. Chicago: Rand McNally and Company.
- Gilbert, Jerome H. 1964. Telsa school breaks the lock step. Elementary School Journal 64 (March):306-309.
- Goldberg, Miriam I.; Passow, A. Harry; and Justman, Joseph. 1966. The effects of ability grouping. New York: Teachers College Press.
- Goodlad, John I., and Anderson, Robert H. 1963. The non-graded elementary school. New York: Harcourt, Brace and World, Inc.
- , and Anderson, Robert H. 1965. Cooperative teaching in educational reform. National Elementary Principal 44 (January):8-13.
- Harvey, Sidney Bilt. 1974. A comparison of kindergarten children in multigrade and traditional settings on self-concept, social-emotional development, readiness development, and achievement (Doctoral dissertation, Virginia Polytechnic Institute and State University). University Microfilms International 35:3340-A.
- Halliwell, Joseph W. 1963. A comparison of pupil achievement in graded and nongraded primary classrooms. The Journal of Experimental Education 32 (Fall):57-64.
- Heathers, Glen. 1969. Grouping. In The grouping of pupils: Thirty-fifth yearbook of the National Society for the Study of Education, Part I. Edited by G. M. Whipple. Bloomington, Ill.: Public School Publishing Company.
- Heddens, James W. 1982. Elementary education. In Encyclopedia of educational research. 5th ed. New York: The Free Press.
- Hillson, Maurie, ed. 1965. Change and innovation in elementary school organization. New York: Holt, Rinehart and Winston.

- , et al. 1964. Controlled experiment evaluating the effects of a non-graded organization on pupil achievement. Journal of Educational Research 57 (July):548-50.
- Hull, J. H. 1958. Multi-grade teaching. Nation's Schools 62 (July):33-36.
- Hollifield, J. H. 1980. Research clues: Do multiage classrooms improve student achievement and self-concept? National Education Association Journal, Elementary Edition 69 (November-December), 14E.
- Isaac, Stephen, and Michael, William B. 1972. Handbook in research and evaluation. San Diego, Calif.: Robert R. Knapp.
- Killough, Charles Kyle. 1971. An analysis of the longitudinal effects that a nongraded elementary program, conducted in an open-space school, had on the cognitive achievement of pupils (Doctoral dissertation, University of Houston). Dissertation Abstracts International 32:3614A.
- Knight, Elton E. 1938. A study of double grades in New Haven city schools. Journal of Experimental Education 7:11-18.
- Koontz, William F. 1961. A study of achievement as a function of homogeneous grouping. The Journal of Experimental Education 30 (December):249-53.
- Lincoln, Robert Douglas. 1981. The effect of single-grade and multi-grade primary school classroom on reading achievement of children. Unpublished doctoral dissertation, University of Connecticut.
- Lunn, Joan C. Barker. 1970. Streaming in the primary school. London: National Foundation for Educational Research in England and Wales.
- MacDonald, P. A., and Wurster, S. R. 1974. Multiple grade primary versus segregated first grade: Effects on reading achievement. Bethesda, Md.: ERIC Document Reproduction Service, ED 094 336.
- Martin, Lyn S., and Pavan, Barbara N. 1976. Current research on open space, nongrading, vertical grouping, and team teaching. Phi Delta Kappan 57 (January): 310-15.

- Milburn, Dennis. 1981. A study of multi-age or family-grouped classrooms. Phi Delta Kappan 64 (March):306-309.
- Miller, Richard I. 1967. The nongraded school: Analysis and study. New York: Harper and Row.
- Morgenstern, Anne. 1963. A comparison of the effects of heterogeneous and homogeneous (ability) grouping on the academic achievement and personal-social adjustment of selected sixth grade children (Doctoral dissertation, New York University). Dissertation Abstracts International 24:1054A.
- , ed. 1966. Grouping in the elementary school. New York: Pitman Publishing Corporation.
- Morris, Vernon Raymond. 1969. An evaluation of pupil achievement in a nongraded primary plan after three and also five years of instruction (Doctoral dissertation). Dissertation Abstracts International 29:3809A.
- Myers, Donald A., and Sinclair, Robert. 1973. Improved decision making for school organization: What and what for. National Elementary School Principal 52 (January):43-50.
- National Education Association, Research Division. 1968. Ability grouping, research summary. Washington, D.C.: National Education Association.
- Otto, Henry J. 1934. Historical sketches of administrative innovations. Educational Administration and Supervision 20 (March):161-72.
- , and Sanders, David C. 1964. Elementary school organization and administration. New York: Appleton-Century-Crofts.
- Pavan, Barbara Nelson. 1973. Good news: Research on the nongraded elementary school. Elementary School Journal 73 (March):333-40.
- Polkinghorne, Ada R. 1950. Grouping children in the primary grades. Elementary School Journal 50 (May):502-508.

- Rankin, P; Andersen, C. T.; and Bergman, W. G. 1936. Ability grouping in the Detroit individualization experiment. In The grouping of pupils: Thirty-fifth yearbook of the National Society for the Study of Education, Part I. Edited by G. M. Whipple. Bloomington, Ill.: Public School Publishing Company.
- Rehwoldt, Walter, and Hamilton, Warren. 1957. By their differences they learn. The National Elementary Principal 37 (December):27-29.
- Reisner, Edward H. 1936. Social changes and its effect on school organization in the United States. In The grouping of pupils: Thirty-fifth yearbook of the National Society for the Study of Education, Part I. Edited by G. M. Whipple. Bloomington, Ill.: Public School Publishing Company.
- Remacle, Leo Felix. 1971. A comparative study of the differences in attitudes, self-concept, and achievement of children in graded and nongraded elementary schools (Doctoral dissertation, University of South Dakota). University Abstracts International 31:5948A.
- Rosenbaum, James E. 1980. Social implications of educational grouping. In Review of Research in Education. Vol. 8. Washington, D.C.: American Educational Research Association.
- Shane, Harold G. 1960. Grouping in the elementary school. Phi Delta Kappan 41 (April):313-19.
- Shearer, William J. 1899. The grading of schools. New York: The H. P. Smith Publishing Company.
- Siegel, Sidney. 1956. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill Book Company, Inc.
- Stendler, Celia B. 1960. Grouping practices. The National Elementary Principal 40 (September):147-65.
- Turney, Austin H. 1936. The psychological basis of grouping. In The grouping of pupils: Thirty-fifth yearbook of the National Society for the Study of Education, Part I. Edited by G. M. Whipple. Bloomington, Ill.: Public School Publishing Company.

- Ward, Dayton Nissler. 1970. An evaluation of a nongraded school program in grades one and two (Doctoral dissertation, The University of Texas at Austin). Dissertation Abstracts International 30:2786A.
- Way, Joyce W. 1969. The effects of multiage grouping on achievement and self-concept. Cortland: State University of New York, Cortland College, Institute for Experimentation in Teacher Education.
- Whipple, G. M., ed. 1936. The grouping of pupils: Thirty-fifth yearbook of the National Society for the Study of Education, Part I. Bloomington, Ill.: Public School Publishing Company.
- Wilhelms, Fred T., and Westby-Gibson, Dorothy. 1961. Grouping research offers leads. Educational Leadership 18 (April):410-13.
- Wilt, Hiram Jack. 1971. A comparison of student attitudes toward school, academic achievement, internal structures and procedures: The nongraded school vs. the graded school (Doctoral dissertation, University of Missouri-Columbia). University Microfilms International 31:5105A.
- Wolfson, Bernise J. 1967. The promise of multiage grouping for individualized instruction. The Elementary School Journal 67 (April):354-56.
- Yates, Alfred, ed. 1966. Grouping in education. New York: John Wiley & Sons, Inc.
- Yerry, Marie J., and Henderson, Edward. 1964. Effects of interage grouping on achievement and behavior (End of Year Report, Experimental Program No. A-27-63). Bethpage, N.Y.: Plainedge Public Schools.